Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 34, NUMBER 4 APRIL, 1944

Physical Concepts: Radiant Energy and Its Measurement*

COMMITTEE ON COLORIMETRY

THE processes of radiation, propagation, transmission, reflection, absorption, and detection of radiant energy are investigated by the use of physical methods and principles, and the concepts and units in terms of which such phenomena are described are physical. The nature of radiant energy and the processes by which it is radiated, modified, and detected have been described in a general manner in Chapter II. The present chapter is concerned with: (1) the definitions of the quantities and units employed for the specification of radiant energy in its various manifestations, and the capacities of materials and objects to radiate, modify, and detect radiant energy; and (2) the fundamental principles and description of suitable apparatus and methods for the measurement of such quantities.

1. RADIANT FLUX AND RADIANT INTENSITY

Radiant energy U, like all other forms of energy, may be measured and specified quantitatively in terms of the erg, the c.g.s. unit of energy. Since the erg is a very small unit, the m.k.s. unit, the joule (which equals ten million ergs) is frequently used. The evaluation of radiant energy, as such, in terms of ergs or joules is appropriate only for the specification of the total energy radiated, or contained in a given volume of space (*radiant density u*), or received in any specific instance. The evaluation of accumulated radiant energy is significant when an accumulator type of detector is used, such as a photographic film or the photoelectric iconoscope employed in television transmitters. For many purposes radiant energy is evaluated in terms of rate of transfer, or *radiant power P*, rather than in terms of the total energy transferred.

Radiant power is usually called *radiant flux*, symbol *P*. It may be specified in terms of ergs per second or watts (joules per second), either the total for all wave-lengths, or per millimicron wave-length interval (per millimicron, in sequel, for brevity) for each spectral region of interest.

2. SPECTRAL COMPOSITION AND SPECTRAL DISTRIBUTION

The *spectral composition* of radiant energy U_λ (or flux P_λ), may be specified by ergs (or watts) per millimicron at each wave-length. The spectral composition may be expressed in terms of whichever spatial variant of radiant energy (or flux) is most appropriate in any specific case, but in this report the term spectral composition will be reserved for data expressed in absolute units. The relative distribution of radiant energy (or flux) is usually sufficient for the solution of a colorimetric problem, and such relative data are much more frequently available than absolute data. The term *spectral distribution* will refer exclusively to relative data specifying the distribution of radiant energy (or flux) at various

* Chapter V of the forthcoming Report of the Committee on Colorimetry. The Committee wishes to acknowledge with appreciation the extensive assistance rendered by Dr. Fred H. Perrin in the preparation of this material.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

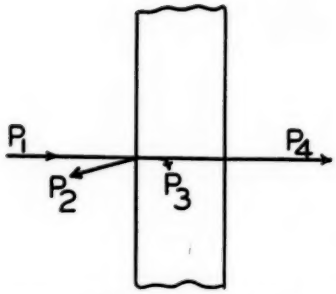


FIG. 1. Schematic diagram indicating incident, reflected, and transmitted beams.

wave-lengths of the spectrum. Spectral distribution data may be supplemented in specific cases by separate statements of the radiant energy (or flux) per unit wave-length interval at some one wave-length, or integrated for all wave-lengths.

3. SPATIAL VARIANTS FOR THE SPECIFICATION OF RADIANT ENERGY

The radiant flux emitted per unit area of a source is the *radiant emittance W*, which may be expressed in terms of the m.k.s. unit, watts per square meter, or in terms of the c.g.s. unit, ergs per second per square centimeter. The flux radiated within a unit solid angle about a source is the *radiant intensity J* of the source, and may be expressed in terms of watts per unit solid angle. This quantity can be measured accurately only by observations made at a distance from the source much greater than the largest dimension of the source; otherwise the solid angle around the source is indefinite. In general, the radiant intensity of a source depends upon the direction from which it is observed. The *average intensity*, however, may be defined as $\frac{1}{4}\pi$ times the total flux radiated by a source.

The radiant flux incident per unit area of a surface is the *irradiance H* of the surface, and may be specified in terms of the m.k.s. unit, watts per square meter, or in terms of the c.g.s. unit, ergs per second per square centimeter. If the irradiated surface is perpendicular to the incident rays, and is at a distance from the source much greater than the largest dimension of the source, and if no absorption or other loss of radiant energy occurs in the intervening space, then the irradiance is equal to the radiant intensity of the source divided by the square of the distance from the source. If the radiant intensity of the source is specified in terms of watts per unit solid angle, and if the distance is expressed in meters, the irradiance calculated in this manner will be expressed in watts per square meter.

Sources of radiant energy frequently have dimensions that are not small in comparison with their distances from the receiver. Radiant intensity is an incomplete and sometimes quite inappropriate specification of such sources. The flux radiated by a unit projected area of the source into a unit solid angle is the *radiance N* of the source. The radiance may therefore be considered as a generalization of radiant intensity expressed per unit projected area of the source. The projected area of the element of the source on which the solid angle is centered, and for which the radiance is specified, is the area of the projection of the surface element onto a plane located at the source and perpendicular to the direction in which the angular flux density is measured. Radiance may be specified in terms of the m.k.s. unit, watts per unit solid angle per square meter, or in terms of the c.g.s. unit, ergs per second per unit solid angle per square centimeter.

4. REFLECTANCE, TRANSMITTANCE, ABSORPTANCE

In the most general case, when radiant flux P_1 falls upon an object (see Fig. 1) some of it is reflected, P_2, from the surface, some of it is absorbed, P_3, and the remainder, P_4, is transmitted, with or without scattering.

Radiant reflectance r is the ratio (P_2/P_1) of the reflected radiant flux to the incident flux. Radiant reflectance for spectrally pure energy depends upon the wave-length of the incident flux, and the value of this quantity at any given wave-length is called the *spectral reflectance r*. No ambiguity results from the omission of the adjective "radiant" in this context, and in fact the unmodified term *reflectance* refers to radiant reflectance in most contexts.

Radiant transmittance t is the ratio (P_4/P_1) of the transmitted radiant flux to the incident radiant flux. *Spectral transmittance* t_λ is the radiant transmittance evaluated for a particular wave-length of incident energy. *Transmittance* refers to

radiant transmittance in most contexts, and exceptions to this rule should be indicated clearly.

Radiant absorptance A is the ratio (P_3/P_1) of the absorbed radiant flux to the incident flux. *Spectral absorptance* A_λ is the radiant absorptance evaluated for a particular wave-length λ of the incident flux. Computation of absorptance as unity minus the sum of reflectance and transmittance is incorrect in many cases because this quantity includes losses by internal reflection of light scattered in the medium, much of which is lost or absorbed at the edges of the object. Similarly, it is usually incorrect to consider the absorptance of an object observed by transmission as the difference between unity and the transmittance. This quantity includes the losses by reflection as well as the losses within the object resulting from both absorption and scattering. In the case of objects observed by reflection, the absorptance cannot be assumed to equal the reflectance subtracted from unity, for this difference includes the transmitted energy as well as the energy that is internally scattered and absorbed.

5. IMPORTANCE OF SPATIAL CONDITIONS OF MEASURING REFLECTANCE AND TRANSMITTANCE

For most surfaces, the measured radiant reflectance, whether spectral or integrated for all of the wave-lengths included in the incident or collected flux, depends upon the spatial distribution of the incident and collected flux. Various instruments employ different angular modes of irradiation and from this cause alone would measure different reflectances. Furthermore, most instruments fail to collect all of the reflected flux and therefore, in general, fail to measure reflectance according to the strict definition of the term; but the quantities measured by such instruments are more or less related to reflectance and the results are loosely referred to as reflectance measurements. An adequate statement of the mode of irradiation and collection is therefore necessary as a supplement to every specification of reflectance.[501–503] When a standard instrument

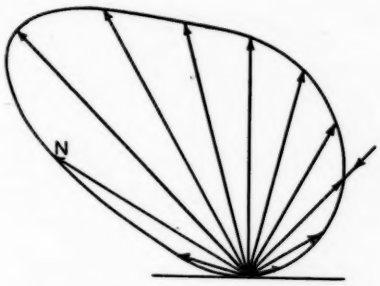


FIG. 2. Goniophotometric curve showing radiance for various directions of collection.

or an apparatus which has been fully described in available publications is employed, the name of the instrument or a reference to the published description is usually sufficient for the specification of the modes of irradiation and collection. *Specular reflectance* is a specialized variety of reflectance, which measures the reflectance of surfaces for image-forming rays. In measurements of specular reflectance the incident flux should be confined near a single direction, within quite narrow angular limits, and only the flux reflected in the corresponding directions of mirror-like reflection should be collected. For surfaces which appreciably scatter the reflected flux, the definition and measurements of specular reflectance are rather indefinite and unsatisfactory, being excessively dependent upon very slight variations in the conditions of measurement.

Measurements of reflectance employing several modes of irradiation and collection are involved in the evaluation of gloss.[504] Reflectances determined with two or more sets of conditions are required for any specification of gloss, and various gloss meters employ different spatial distributions of incident and collected flux. Consequently, such instruments imply different definitions of gloss, and frequently result in quantitatively different specifications of gloss. Goniophotometers measure reflectances with collection confined to a narrow range of angles, the central ray of which is variable over a wide range of angles relative to the reflecting surface. Figure 2 is a typical goniophotometric curve showing by the lengths of the lines N, the

[501] H. J. McNicholas, Bur. Stand. J. Research **1**, 29–73 (1928).
[502] U. S. National Bureau of Standards, "The reflectance of paints and pigments," Letter Circular, LC-470 (1936).
[503] K. S. Gibson, Paper Trade J. (technical section) **111**, 135–139 (1940).
[504] R. S. Hunter, J. Research Nat. Bur. Stand. **18**, 19–39 (1937).

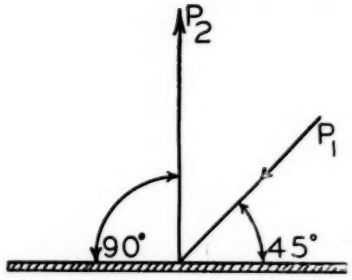


FIG. 3. Standard conditions for irradiation and collection.

radiance of a paper for the corresponding angles of collection. Various conditions of irradiation are used in different goniophotometers, and these conditions should be specified in the reports of the data obtained with any specific instrument. In some instruments the incident energy is almost unidirectional, from either a fixed or variable direction. In other instruments the incident energy is more or less completely diffused. Goniospectrophotometers employ spectrally isolated energy for reflectance measurements.

The International Commission on Illumination has recommended standard conditions of irradiation and collection, for use in reflectometry.[505] These conditions (see Fig. 3) call for unidirectional irradiation at an angle of 45 degrees from the normal to the reflecting surface, and unidirectional collection of the flux reflected along the normal. No specification of the tolerance in these angles has been adopted, although all instruments require incident and collected beams varying by finite angles from the central rays. Considerable departures from the standard conditions recommended by the I.C.I. are frequent in practical reflectometry.[503] Diffuse reflectance is measured by instruments in which either the incident or the collected flux is distributed approximately uniformly among all angles between the normal and the plane of the reflecting surface. The latter case is illustrated in Fig. 4, in which an integrating sphere is used for collection of the reflected flux. According to the Helmholtz reciprocal relationship, the spatial distributions of the incident and the reflected flux may be completely interchanged without changing the measured reflectance. By virtue of this rule many instruments which employ quite different optical devices measure very similar varieties of reflectance.

As in the case of reflectance, the measured transmittance of an object depends upon the spatial distributions of incident and collected flux. Precise descriptions of the modes of irradiation and collection are necessary to specify transmittance completely. *Regular transmittance* measures the fraction of the incident flux that is transmitted without scattering. In the presence of appreciable scattering, regular transmittance cannot be defined or measured precisely.

6. REFLECTION: LAMBERT'S LAW AND FRESNEL'S FORMULA

When a beam of radiant energy is incident on the boundary between media of different refractive index, a certain amount is reflected. In case the boundary is plane and polished, the reflected beam leaves the boundary at an angle from the normal equal to the incident angle. This case is referred to as *specular* (mirror-like) *reflection*. In case the boundary is rough, specular reflection can be considered to take place at a great number of very small plane areas oriented in various directions, as in Fig. 5, so that the reflected energy is distributed over a wide range of angles. *Perfectly diffuse reflection* is defined as that described by *Lambert's law:* The flux

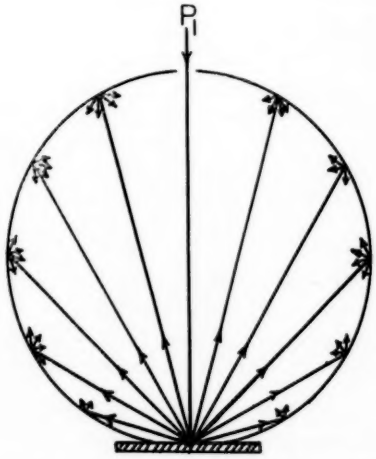


FIG. 4. Integrating sphere for measurement of diffuse reflectance.

[505] International Commission on Illumination, *Proceedings of the Eighth Session* (Cambridge, 1931), pp. 19–29.

reflected per unit solid angle is proportional to the cosine of the angle measured from the normal to the surface. Since the projected area varies inversely as the cosine of this angle, the radiance of a perfect diffuser is independent of the viewing angle.

The cases just described are ideal; reflection at a real surface is always partly specular and partly diffuse, and in general the specification of the behavior of reflecting surfaces requires the measurement and statement of reflectance for several conditions of irradiation and collection.

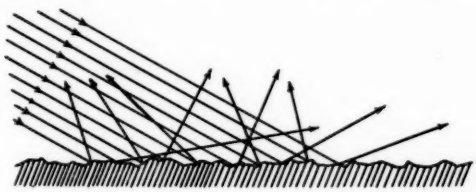


FIG. 5. Schematic diagram of reflection from opaque diffusing surface.

The reflectances of transparent dielectrics with plane polished surfaces are specified for normal incidence by Fresnel's formula:

$$r_f=(n'-n)^2/(n'+n)^2. \qquad (1)$$

In this formula n represents the index of refraction of the medium (usually air) surrounding the reflecting object, and n' represents the index of refraction of the reflecting medium. If the incident beam is unpolarized, and if the initial medium has the lower index, the reflectance as indicated in Fig. 6 rises as the angle of incidence is increased, gradually at first and then rapidly until it becomes unity at 90 degrees. Tables specifying Fresnel reflectances for various angles, and for both polarized and unpolarized incident energy, have been published by Moon.[506] The remainder of the incident energy enters the reflecting medium, at an angle of refraction i', according to Snell's law, where i represents the angle of incidence:

$$n \sin i=n' \sin i'. \qquad (2)$$

When the initial medium has the higher index, a similar behavior is observed for angles less than a certain critical angle, for which $\sin i=n'/n$.

[506] P. Moon, J. Math. and Phys. **19**, 1–33 (1940).

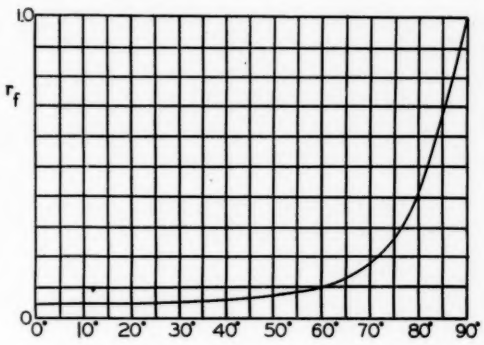


FIG. 6. Dependence of Fresnel reflectance on angle of incidence.

For greater incident angles, the beam is totally reflected into the initial medium.

This *Fresnel reflection*, as it is frequently termed, is of great importance in practical optics. Thus, glass of index 1.5 has a Fresnel normal reflectance of 4 percent at each surface, so that the total loss of transmitted light by reflection at the surfaces of a simple lens is $(1.00-0.96\times0.96)$, or approximately 8 percent. In a multilens system this loss becomes serious, not only because it weakens the image but also because the reflected energy appears as ghosts, flare spots, and veiling glare. On the other hand, the phenomenon of total reflection is employed with advantage in refractometers and other instruments using totally reflecting prisms. The critical angle for glass of index 1.5 is 41 degrees, so that

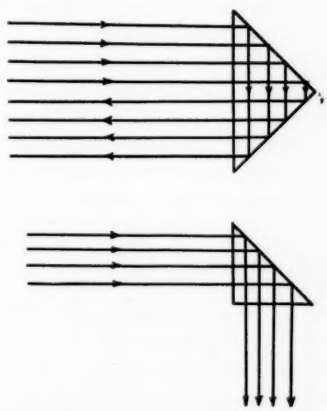


FIG. 7. Totally reflecting prisms for reversing direction of beam and rotating beam through a right angle.

a 90-degree isosceles prism arranged as in Fig. 7 will reverse the direction of a beam by two internal reflections, or turn it through 90 degrees by a single reflection at the hypotenuse. Such prisms are employed in prism binoculars and in many other optical instruments.

7. TRANSMISSION AND ABSORPTION: BOUGUER'S AND BEER'S LAWS

Of the radiant energy that traverses even a transparent material, a certain fraction is absorbed, and it is reasonable to expect that equal layers will absorb equal fractions of each kind of energy entering them. This was pointed out by Bouguer in 1729, whence it is known as *Bouguer's* (sometimes incorrectly as Lambert's) *law*. If this law is applied to infinitesimal layers and the losses integrated, the radiant flux at a distance x within the medium is found to be

$$P = P_0 e^{-kx}, \tag{3}$$

where P_0 is the flux entering the first layer and k is a constant, known as the *absorption coefficient*. Most substances absorb selectively, which merely means that k is a function of wave-length λ. Such substances are said to be *selective*. In some cases k_λ varies so rapidly that, with heterogeneous irradiation, thin layers have one color and thick layers quite another. Thus, in some materials which are blue in thin layers, k_λ is so much greater for all of the short waves than for the extreme long waves of the visible spectrum that thick layers of the same material appear deep red. This phenomenon is known as *dichroism*, and materials exhibiting this effect are said to be *dichroic*.

Bouguer's law is frequently applied to solutions, in which case it would seem to be a matter of indifference whether the concentration or the thickness is varied. Beer drew attention to this rule in 1852, and it is confirmed by observation whenever the process of dilution does not affect the solute or solvent and the latter is either non-absorbing or a suitable correction is made. If c represents the concentration, *Beer's law* states that the radiant flux transmitted by a layer of thickness x is

$$P = P_0 \exp(-\beta_\lambda c x), \tag{4}$$

where β_λ is the *specific absorption coefficient*. When the concentrations are expressed in gram-molecular weights, and x is measured in centimeters, β_λ is known as the *molecular extinction coefficient*.

Another quantity sometimes used to express the absorption characteristics of a substance is the *extinction coefficient*, commonly denoted by κ. This is defined by the statement that the amplitude of a wave of length λ when traversing a layer of the substances having a thickness λ decreases in the ratio $1 : e^{-2\pi\kappa}$. Since the radiant flux is proportional to the square of the amplitude, the flux at a distance x within the medium is

$$P = P_0 e^{-4\pi\kappa x/\lambda}. \tag{5}$$

The distances x and λ should be expressed in the same units for this calculation. The coefficients of x in Eqs. (3) and (5) may be equated, with the result:

$$\kappa = k\lambda/4\pi. \tag{6}$$

8. TRANSMITTANCE, OPACITY, AND DENSITY

The *radiant transmittance* t of an object has been defined as the transmitted flux P_t divided by the incident flux P_i;

$$t = P_t/P_i. \tag{7}$$

The radiant transmittance for any specific wavelength is called the *spectral transmittance* t_λ. The reciprocal of transmittance is called opacity,

$$o = 1/t = P_i/P_t. \tag{8}$$

This use of the term "opacity" should not be confused with the usage which defines opacity in terms of the contrast ratio of reflectances of semi-opaque materials, such as paper, when backed successively with a black-lined cavity and with a surface of magnesium oxide.[507–511]

---

[507] M. N. Davis, "Methods and instruments for opacity measurement," Tech. Assn. (Pulp and Paper Ind.) Papers **15**, 347–350 (1932).

[508] *Official Method for the Determination of the Opacity of Paper* (Technical Association of the Pulp and Paper Industry, New York).

[509] D. B. Judd, Bur. Stand. J. Research **12**, 345–351 (1934).

[510] D. B. Judd, J. Research Nat. Bur. Stand. **13**, 281–291 (1934).

[511] U. S. National Bureau of Standards, "Standards for checking opacity meters," Letter Circular, LC-418 (1934).

The term *density*, common in photography and defined as

$$D = \log o = \log (1/t) = \log (P_i/P_t), \qquad (9)$$

is also widely used in other applications of optics. If compensations for surface reflectances are made, and if Beer's law is obeyed, the spectral density of a solution is proportional to the thickness and the concentration. If interreflections between the several surfaces are eliminated, the spectral density of a pile of transparent non-diffusing plates equals the sum of the spectral densities of the individual plates. In general, the densities of selective ("colored") materials, when measured for radiant energy distributed throughout a broad wave-length band, are not additive and do not obey Beer's law.

9. INTERNAL TRANSMITTANCE, TRANSMISSIVITY, ABSORPTIVITY

In the presence of interreflections the convenient rule of the additivity of densities is only approximate. The simplest case, i.e., the transmittance of a pile of non-absorbing, non-diffusing plates was investigated by Stokes[512] in 1860, but the general solution, including the effects of absorption, is not yet completed and each specific case must be investigated separately. As a result of Fresnel reflections at the surfaces and interreflections between the surfaces, the spectral transmittances of single plates of varying thicknesses of a given material do not obey Bouguer's law, but quite accurate approximations can be computed for the magnitudes of the effects of surface reflections and interreflections in any specific instance. The *internal transmittance* t_i of an object is the ratio of the flux, P_3 in Fig. 8, incident on the second surface to the flux P_2, transmitted by the first surface. It can usually be computed with satisfactory accuracy from the measured transmittance t (the ratio P_4/P_1) by multiplying by a factor which is dependent on wave-length. If the Fresnel reflectances, $r_f = (P_1 - P_2)/P_1$, $r_f' = (P_3 - P_4)/P_3$, are equal for the two surfaces, then to a close approximation, ignoring the effects of interreflections between the two surfaces,

$$t_i = t/(1 - r_f)^2. \qquad (10)$$

The value of the internal transmittance for a unit thickness of a non-diffusing substance is the *transmissivity*. The transmissivity subtracted from unity is the *absorptivity* of the substance. For spectrally pure radiant energy, the internal transmittance, $t_{i\lambda} = P_3/P_2$, can be computed from Bouguer's and Beer's laws, to the accuracy with which $t_{i\lambda}$ can be determined for media obeying those laws:

$$t_i = \exp(-k_\lambda x) = \exp(-\beta_\lambda c x). \qquad (11)$$

The *internal density* for the corresponding wave-length can be defined as the logarithm of the reciprocal of the *spectral internal transmittance* $t_{i\lambda}$, and this density for such media is proportional to thickness and concentration, or the product of these:

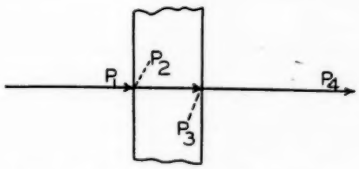


FIG. 8. Schematic diagram indicating losses by external and internal surface reflections.

$$D_{i\lambda} = \log_{10}(1/t_{i\lambda}) = mk_\lambda x = m\beta_\lambda c x, \qquad (12)$$

where $m = \log_{10} e$.

If the absorptance of the solvent is appreciable, internal density is not proportional to concentration of the solution and frequent use is made of *transmittancy:* the ratio of the transmittance of the solution to that of the solvent in equivalent thickness. In the case of solutions which obey Beer's law, the logarithm of transmittancy, regardless of the absorptance of the solvent, is proportional to concentration.

10. TRANSMITTANCE AND REFLECTANCE OF TURBID MEDIA

In the preceding sections, it has been assumed that the media are *homogeneous* and *isotropic*. All infinitesimal elements of volume are identical in composition and optical properties in a *homogeneous* material; the velocity of propagation is identical for all directions in an *isotropic*

[512] G. G. Stokes, Proc. Roy. Soc. **11**, 545 (1860).

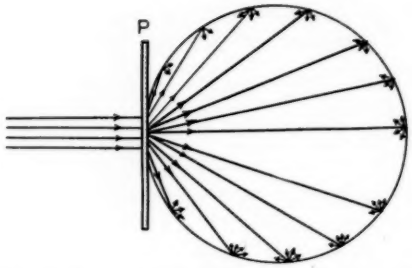


FIG. 9a. Measurement of diffuse transmittance by collection in integrating sphere.

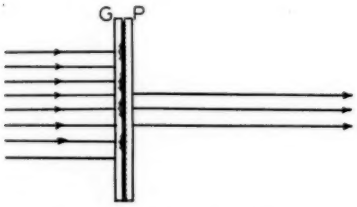


FIG. 9b. Measurement of diffuse transmittance by opal glass irradiation and unidirectional collection.

material. The velocity depends upon the direction of propagation in *anisotropic* or *aeolotropic* media. *Heterogeneous*, or *turbid*, media exhibit fine-grained variations in composition or optical properties which cause appreciable scattering of radiant energy.

When the heterogeneities have average dimensions somewhat smaller than the wave-length of the incident energy, the energy is scattered from the direction of rectilinear propagation, and the fraction of the incident flux thus scattered is inversely proportional to the fourth power of the wave-length. This is known as *Rayleigh's law of scattering*, and the phenomenon, which is responsible for the blue of the sky and the iris of the human eye, is called the *Tyndall effect*. When the heterogeneity is on a larger scale, the scattering is non-selective. A coarse-grained photographic deposit is an example for wave-lengths in the visible portions of the spectrum, the dimensions of the silver grains being of the order of 2μ and more. Intermediate cases, exhibiting less selectivity than the Tyndall effect, correspond to heterogeneities of the same order of magnitude as the wave-length of the incident flux.

The flux transmitted by any turbid medium emerges with a very wide angular distribution, and a statement of the spatial distributions of both the incident and collected beams is necessary to supplement values of transmittance obtained with any particular apparatus. The angular comprehension of an instrument should correspond to the conditions under which the measured materials are to be used. Thus, when a photographic plate is used for contact printing, practically the entire emergent beam is effective, and a suitable photometer would be one measuring the amount of flux entering an integrating sphere through a small hole covered by the plate *P*, as indicated in Fig. 9a. It is customary, however, to reverse the arrangement as shown in Fig. 9b by placing the emulsion side of the photographic plate in contact with an opal glass plate *G* which diffuses the incident beam.[513] The ratio of the radiance of the photographic plate thus irradiated to the radiance of the opal glass alone is called the *diffuse transmittance* of the plate. The logarithm of the reciprocal of the diffuse transmittance measured under these conditions is called the *diffuse density*. The proportion of the flux of a collimated beam transmitted without deviation by the photographic plate is called the *specular transmittance*, and is lower than the diffuse transmittance. The logarithm of the reciprocal of specular transmittance is called *specular density*. The ratio between the specular and diffuse density of a medium is called the *Callier coefficient*, and is denoted by *Q*. It always exceeds unity, and is of the order of 1.4 for ordinary photographic deposits.

Neither the transmittances nor the densities of turbid media obey any simple formulas, such as Bouguer's or Beer's laws. The optical behaviors of such media have been the subject of very extensive and highly complicated investigations,[514–518] the results of which are very important in some applications of colorimetry.

---

[513] A. M. Koerner and C. Tuttle, J. Opt. Soc. Am. **27**, 241–256 (1937).
[514] H. J. Channon, F. F. Renwick, and B. V. Storr, Proc. Roy. Soc. **A94**, 222–237 (1918).
[515] J. W. Ryde, Proc. Roy. Soc. **A131**, 451–464 (1931).
[516] J. W. Ryde and B. S. Cooper, Proc. Roy. Soc. **A131**, 464–475 (1931).
[517] P. Kubelka and F. Munk, Zeits. f. tech. Physik **12**, 593–601 (1931).
[518] S. Q. Duntley, J. Opt. Soc. Am. **32**, 61–70 (1942).

11. INTEGRATION OF SPECTRAL TRANSMITTANCE, REFLECTANCE, AND ABSORPTANCE

Unless a medium is non-selective with respect to wave-length, transmittances, reflectances, densities, and absorptances have no significance in the absence of wave-length specifications. For heterogeneous incident flux some sort of integration must be performed, either explicitly or implicitly. The total radiant flux transmitted by a filter, for example, may be expressed as

$$P = \int_0^\infty P_\lambda t_\lambda d\lambda, \qquad (13)$$

where P_λ specifies the spectral composition of the incident flux and t_λ the spectral transmittance of the filter. The step-by-step determinations of the spectral transmittance, the spectral composition of the incident flux, and the subsequent computations may be avoided in practice by employing a non-selective detector of radiant energy for the measurement of the transmitted flux. A thermopile is such a detector for rather great portions of the spectrum.

The dependence of radiant transmittance t on spectral transmittance t_λ is indicated by the formula:

$$t = \int_0^\infty P_\lambda t_\lambda d\lambda \Big/ \int_0^\infty P_\lambda d\lambda. \qquad (14)$$

The transmittances and the corresponding densities determined for heterogeneous incident flux do not obey either Bouguer's or Beer's laws except in the case of perfectly non-selective media, and densities of selective materials determined with heterogeneous flux are not additive.

12. POLARIZATION

Hitherto we have neglected any consideration of the azimuth of the transverse vibrations of the waves which are associated with radiant energy. Now, in the case of Fresnel reflection from the surface of a refracting substance, if the electric vibrations of the incident waves are parallel to the plane of incidence, and if the angle of incidence is such that the angle between the reflected and refracted beams equals 90 degrees, the entire beam is transmitted, and none is reflected.[506] This angle of incidence is called the *polarizing angle* and its definition is *Brewster's law*. Ordinary radiant energy behaves as though it consisted of many series of waves with planes of vibration oriented at all possible azimuths, and for this reason it is considered *unpolarized*. When unpolarized radiant energy is incident at the polarizing angle, as in Fig. 10, the reflected beam is completely polarized with a single plane of vibration. The *plane of polarization* was defined (long before the invention of the electromagnetic theory of radiant energy) as the plane of incidence (and reflection) of the beam reflected at the polarizing angle. According to the electromagnetic theory, the oscillating magnetic field $\mathfrak{H}$ of a plane-polarized beam is parallel to the plane of polarization, and the oscillating electric field $\mathcal{E}$ of such a beam is always perpendicular to the plane of polarization. Because the alternating electric field, rather than the magnetic field, seems more representative of the transverse vibrations of radiant energy, the plane parallel to the electric field (thus perpendicular to the plane of polarization) is called the *plane of vibration*.

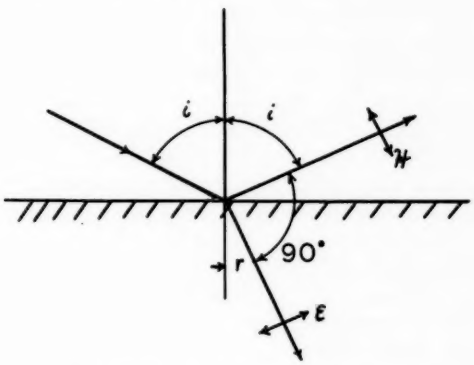


FIG. 10. Brewster's law and definition of plane of polarization.

The subject of polarization, and of crystal optics with which it is intimately associated, is much too complex to be covered in this brief treatment. Polarization phenomena are widely applied in photometry and spectrophotometry, however, and precautions are necessary to avoid errors arising from the use of polarizing instruments for the measurement of substances which partially or completely polarize radiant energy. A short general discussion of the properties of

polarizing substances is therefore appropriate in an extensive discussion of colorimetry.

Polarizing substances are usually crystalline and have different indices of refraction for different directions of passage through the crystals. An unpolarized ray entering such a crystal is separated into two plane polarized rays, having perpendicular planes of polarization related to the crystalline axes as well as the direction of the separated rays. In general the two components travel with different velocities. In *uniaxial* crystals there is a single direction along which the two rays are propagated with the same velocity. In *biaxial* crystals there are two directions of *single ray velocity*. The optical behavior of a crystal is determined by the crystallographic system to which it belongs. Those belonging to the cubic system (e.g., fluorite, sylvine, rocksalt) are isotropic and show no polarizing effects. Those belonging to the tetragonal and the hexagonal systems (e.g., quartz, calcite) are uniaxial, and crystals belonging to other systems are biaxial. Biaxial crystals are not commonly used in optical instruments and need not be considered here in further detail.

A ray incident on a unixial crystal is separated into plane polarized rays, one of which (called the ordinary ray) obeys Snell's law and behaves in every respect as if the crystal were isotropic. The other (called extraordinary) ray does not obey Snell's law, as is evident in Fig. 11, and the refractive index depends on the direction of the ray in the crystal. In a *positive* or *prolate* crystal such as quartz, the index of refraction of the extraordinary ray is greater than the index for the ordinary ray. In a *negative* or *oblate* crystal, such as calcite, the index of the extraordinary ray is less than that of the ordinary ray. Crystals with appreciably different indices for the ordinary and extraordinary rays are said to be *doubly refracting*. Lenses and prisms made of doubly refracting materials should be cut so that the central ray of the utilized path has the same direction as the optic axis, in order to minimize aberrations due to double refraction.

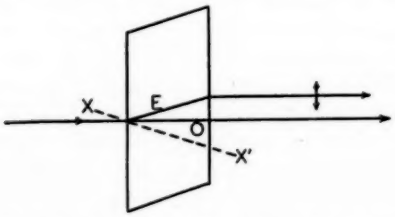


FIG. 11. Double refraction in an uniaxial crystal. Arrow indicates direction of vibration of extraordinary ray.

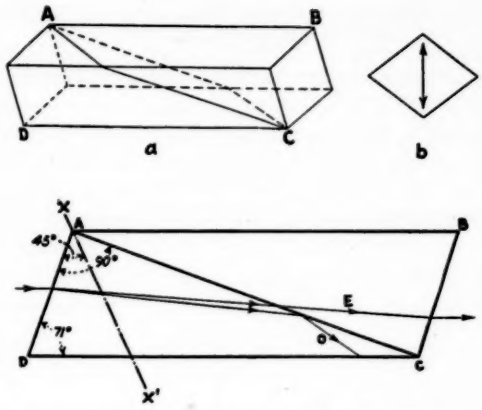


FIG. 12. *Upper:* Diagrammatic sketch of a Nicol prism. The direction of vibration of the emergent beam is indicated in the end elevation at *b*. *Lower:* Cross section of a Nicol prism through the corners *A*, *B*, *C*, and *D* in upper diagram. The values of the angles are those of Nicol's original prism. The optic axis is indicated by *XX'*.

13. POLARIZING PRISMS

The properties of double refracting crystals can be utilized to produce plane-polarized beams. Some natural crystals absorb one of the beams, but most of these crystals are too small to be useful or, as in the case of tourmaline, even the transmitted beam is seriously absorbed in some portions of the spectrum. The synthetic material known as "Polaroid," which consists essentially of minute crystals arranged with their axes parallel and embedded in a transparent sheet, also largely absorbs one beam. For the highest purity and efficiency of polarization, the type of prism devised a hundred years ago by William Nicol is still unexcelled. As indicated in Fig. 12,* a rhomb *ABCD* of calcite is cut along the diagonal *AC*, and after the cut faces are polished they are recemented with Canada balsam. The optic axis is *XX'*. An unpolarized beam entering one end of the rhomb is divided by the double refraction of the calcite, and the

* Reproduced from *Principles of Optics*, A. C. Hardy and F. H. Perrin, by permission of McGraw-Hill Book Company, New York.

extraordinary component E emerges at the other end. It vibrates in the plane of the diagram $ABCD$, which is called the *principal plane* of the prism. The refractive index and consequently the angle of incidence on the balsam layer are both greater for the ordinary than for the extraordinary ray. As a result the ordinary component O is totally reflected at the balsam layer. Many modified forms of the Nicol prism are now in use, a common one being the Glan-Thompson prism which has normal end faces and a greater field angle. In general, the modified polarizing prisms are more difficult to prepare and have less durable surfaces than the original form invented by Nicol.

When two beams polarized in mutually perpendicular azimuths are desired, two wedges of calcite can be cemented together, as shown in Fig. 13,* with their optic axes as indicated in the cross-hatching. In the Rochon prism A the ordinary beam is undeviated and this prism is occasionally used in place of a prism of the Nicol type, for which purpose the deviated extraordinary beam is intercepted and absorbed some distance from the prism. In the Wollaston prism B both beams are deviated equally on opposite sides of the direction of the incident beam.

One of the most common photometric instruments using the Wollaston prism is the Martens photometer. In this instrument a pair of beams from a Wollaston prism passes through a single Nicol prism, which can be rotated by the observer to make the two emergent beams equally intense. If the principal plane of the Nicol prism makes an angle θ with the plane of vibration of one of the beams, the fraction of the amplitude of that beam transmitted by the Nicol prism is proportional to $\cos\theta$ and the fraction of the amplitude of the other beam transmitted is proportional to $\sin\theta$. The relative intensity of the two beams incident on the Nicol prism is therefore $\sin^2\theta/\cos^2\theta = \tan^2\theta$. If the angle θ necessary to secure a balance of the beams transmitted by the Nicol prism is 45 degrees, the two incident beams are equal. When an absorbing material is placed in the path of one beam, the balance of the transmitted beams is disturbed and the square of the tangent of the angle required to restore a balance is equal to the transmittance of the absorbing material. The Martens photometer is discussed further in Section 26.

14. OPTICAL ACTIVITY AND ROTATORY DISPERSION

So far in this discussion, the vibrations have been assumed to take place in a single plane that remains fixed in orientation in each medium. When a plane-polarized beam is passed through solutions of certain substances, however, the plane of the vibration is rotated through an angle that is proportional both to the length of the path of the beam in the solution and to the concentration of the solution. Many solids, liquids, and gases, produce this same phenomenon, which is called *optical activity*. The materials are said to be optically active, and the direction and angle of rotation for a unit path length are characteristic of each substance. Quartz is optically active, but fortunately two varieties having equal and opposite rotatory characteristics occur so that the optical activity can be neutralized in complete optical systems by using equal quantities of right- and left-handed quartz. The central ray of the beam should travel parallel to the optic axis of each quartz element in such a system. Optical activity is approximately proportional to the reciprocal of the square of the wave-length of the radiant energy. The planes of polarization of the spectral components of a beam of plane-polarized heterogeneous radiant energy are therefore separated by optically active media so that the spectral distribution of the beam transmitted by an analyzing Nicol

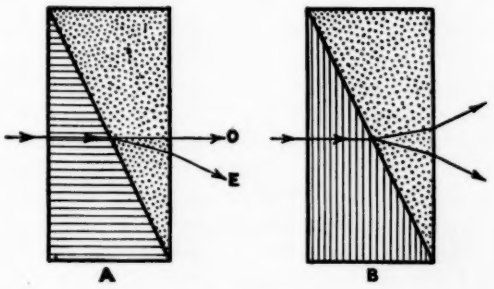


FIG. 13. A, Rochon prism. B, Wollaston prism. The optic axes are indicated by the cross-hatching.

* Reproduced from *Principles of Optics*, A. C. Hardy and F. H. Perrin, by permission of McGraw-Hill Book Company, New York.

prism can be varied by varying the orientation of the latter. This phenomenon is called *rotatory dispersion* and it has been applied by Priest[519] to the production of artificial daylight of continuously variable spectral distribution for photometric comparisons. Such an instrument could be used for a general purpose colorimeter, and Buchwald[520] has worked out in some detail the chromaticities which could be obtained with suitable arrangements of such an instrument. The colors observed by placing thin sections of minerals between Nicol prisms (as in a petrographic microscope) greatly assist the classification and identification of the crystalline constituents.

15. INTERFERENCE OF POLARIZED LIGHT

Wave trains polarized in planes of any mutual orientation can interfere, but we need consider further only trains polarized in perpendicular azimuths because trains having planes of polarization at any other angle can be resolved into perpendicular components. The phenomena of interference are in general detectable only if the interfering beams are coherent (in practice, derived from the same area of a single source). Independent of the relative amplitudes of the interfering waves, if they are exactly in phase or exactly 180 degrees out of phase, their resultant will be another plane-polarized wave, and the plane of polarization of this resultant can be determined by drawing the vector diagram of the vibrations perpendicular to the common direction of propagation. When the phase difference is exactly 90 degrees and the amplitudes are equal, the resultant wave is said to be circularly polarized. In the general case when the amplitudes are not equal or the phase difference of the interfering waves is not 90 degrees, the resultant is elliptically polarized. The quarter-wave retardation required for transforming plane-polarized radiant energy into a circularly polarized beam can be introduced by a suitably oriented lamina of a doubly-refracting crystal cut perpendicular to its optic axis of such a thickness that the optical path lengths of the ordinary and the extraordinary rays differ by exactly one-quarter of a wave-length. Regardless of its orientation, such a *quarter-wave plate* will transform circularly polarized waves into plane-polarized waves, and when suitably oriented a quarter-wave plate will transform elliptically polarized waves into plane-polarized waves. Such a plate may be used with a Nicol prism for the detection and analysis of circular and elliptically polarized beams, but more refined devices which need not be discussed here are necessary for even fair precision in such analyses.

16. ELIMINATION OF POLARIZATION ERRORS IN PHOTOMETRY

Polarized light is more common in nature than is often realized. Skylight, which is scattered sunlight, can be seen to be strongly polarized by examining it with a rotating Nicol prism or a sheet of Polaroid. When Rayleigh derived his fourth-power law of scattering, he showed that the scattered rays must be polarized with their planes of polarization passing through the source, the observer, and the scattering particle. The radiant energy from an incandescent metal surface is polarized to an extent that varies from zero at normal emergence to complete polarization of the energy which is radiated tangent to the surface. An effective way to destroy polarization is to pass the beam through a substance that scatters non-selectively, such as opal glass. Diffusion alone is not sufficient to depolarize; considerable polarization persists after polarized energy is transmitted through a glass plate with a ground surface, or after reflection from a diffusing screen, especially depolished metallic screens or screens painted with metallic paints.

Inasmuch as photometric apparatus frequently employs polarizing elements and depends for accuracy upon geometrical relationships between those elements, the existence of accidental polarization resulting from extraneous causes should never be overlooked as a source of possible error. Polarization can be introduced or modified by reflection at the surface of non-conductors and still more so at the surfaces of conductors. Materials such as cloth with a pronounced luster or sheen will produce noticeable polarization effects. Errors arising from such effects can be eliminated in some photometric instruments by rotating the sample rapidly in its own plane. In

[519] I. G. Priest, J. Opt. Soc. Am. **7**, 1175–1209 (1923).
[520] E. Buchwald, Ann. d. Physik [5] **38**, 245–260 (1940).

other instruments the error is most conveniently avoided by averaging the reflectances observed when the sample is placed alternately in the two beams entering the photometer, maintaining the orientation of the sample unchanged. In general, for all polarizing photometers, the average between the maximum and minimum observed reflectances, or between any two reflectances observed with two orientations of the sample differing by exactly 90 degrees, is the correct reflectance for unpolarized energy. The fact that a quarter-wave plate produces a quarter-wave retardation for only one wave-length prevents the use of such a device for the elimination of the errors arising from the use of polarization pho-

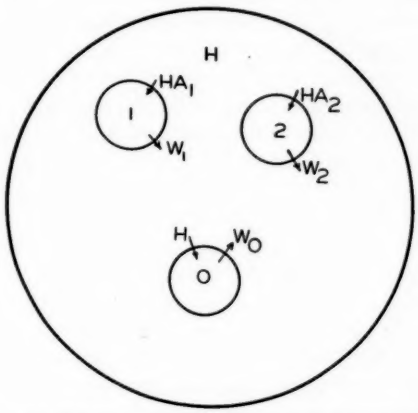


FIG. 14. Schematic diagram of evacuated constant temperature enclosure and radiation equilibrium between several substances.

tometers for the measurement of polarizing materials. The use of circularly polarized light incident on a polarizing material would give the same results as completely unpolarized light, but there is no simple device for the conversion of all of the spectral components of heterogeneous plane-polarized energy into circularly polarized waves.

Unsuspected strains in lenses or prisms, due to mechanical stresses, imperfect annealing, or considerable gradients of temperature, convert plane-polarized waves into elliptically polarized waves, or rotate the plane of polarization. Consequently, most modern photometers contain no optical elements between the polarizer (e.g., the Wollaston prism) and the analyzer (e.g., the Nicol prism). Even the slanting faces of Nicol prisms introduce perceptible amounts of elliptical polarization, so modern modifications of Nicol prisms with faces normal to the central rays of the incident beams have the secondary advantage of eliminating this possible source of error. The polarization as well as the serious losses caused by Fresnel reflections at oblique surfaces of spectroscopic and other prisms can be greatly reduced by treatment of the surfaces to reduce such reflection losses.

Many materials which are ordinarily isotropic become anisotropic or optically active, or both, when subjected to mechanical strain, thermal gradients, and electric or magnetic fields. Such effects should be considered in the use as well as in the design of polarizing instruments, including many spectrophotometers and colorimeters. The influence of mechanical strain on optical properties has an important application in the engineering analysis of the strains in structures, for which purpose models are constructed partially or entirely of such photoelastic materials and loaded to simulate the conditions expected in the contemplated structures.

17. THERMAL RADIATION

Most of the artificial sources of radiant energy encountered in photometry or used in colorimetry consist of incandescent solids. The vibrating particles in such sources are so closely packed that their motions are chaotic and not characteristic of the composition of the particles. The spectrum of radiant energy produced by such motions spreads over all wave-lengths in very extensive spectral regions. Spectral distributions of this sort are called *continuous*.

A fundamental relationship between the radiant emittance of an object and its absorptance can be derived from general thermodynamical principles. Suppose two objects are suspended in an evacuated enclosure, the walls of which are held at some constant temperature. Under such conditions, each object can gain or lose energy only in the form of radiant energy. According to Prevost's law of exchanges, a condition of equilibrium will finally be reached and both objects will reach and remain at the temperature of the walls of the enclosure represented in Fig. 14. In this condition of equilibrium these objects

must radiate energy at the same rate as they absorb energy. If W_1 and W_2 represent the radiant emittances of the two objects, and if A_1 and A_2 represent the corresponding absorptances, the condition of equilibrium requires that

$$W_1/A_1 = W_2/A_2, \qquad (15)$$

which rule may be extended to include any number of objects in the evacuated, constant-temperature enclosure. This is the law derived by Kirchhoff in 1859, which may be stated as follows: The ratio of the radiant emittance to the radiant absorptance for any object is the same as for all other objects at the same temperature.

At any temperature, the radiant emittance of a source is directly proportional to its absorptance, consequently an object for which the absorptance is unity is the best possible radiator. Because it completely absorbs all incident energy and reflects none, such an object is usually called a "blackbody." This name for the ideal radiator is confusing and undesirable because, although it does not reflect any energy, it does radiate, and in most cases of interest in photometry and colorimetry it radiates copiously. Since the use of the name "black" with reference to the most efficient radiators seems so anomalous in a report on colorimetry, the alternative term *complete radiator* is used throughout this report. The radiant emittance from a complete radiator is the maximum that can be attained as a result of temperature alone. For a complete radiator, therefore, the reflectance is equal to zero and the absorptance is equal to unity. If W_0 designates the radiant emittance of a complete radiator, then *Kirchhoff's law* can be written:

$$W_1/A_1 = W_2/A_2 = W_0. \qquad (15a)$$

This law applies also to the emittance and the absorptance for each separate wave-length. No substances have been discovered which completely absorb all the energy incident upon them; no complete radiators occur in nature. A hollow opaque enclosure, in the wall of which is an aperture very small in comparison with the internal dimensions of the enclosure, provides a practical realization of a completely absorbing object, and such a device radiates as a complete radiator. Practically all energy passing through the aperture into the enclosure is reflected and scattered about the interior until it is absorbed, only a negligible proportion escaping through the aperture. When the walls of such an enclosure have uniform temperature, the energy they reflect is just sufficient to supplement the energy they radiate, so that they would be in equilibrium with a complete radiator at the same temperature placed in the enclosure. The small sample of energy radiated outwards through a small aperture does not appreciably disturb this equilibrium, but it is representative of the energy emitted by a complete radiator.

A study of some of Tyndall's experimental observations led Stefan to suggest that the radiant emittance of any surface is proportional to the fourth power of its absolute temperature. Boltzmann, from purely theoretical considerations of a thermodynamic nature, subsequently derived this law for a complete radiator. Boltzmann's reasoning was concerned with the pressure which Maxwell had proved was exerted on a surface by radiant energy incident upon it. The constant of proportionality in the *Stefan-Boltzmann* law has been determined very accurately, and adopting the value recommended by Birge[521] the law is:

$$W = 5.672 \times 10^{-5} T^4, \qquad (16)$$

in which W is expressed in ergs per second per square centimeter, and T is expressed in degrees (Kelvin) absolute temperature.

18. SPECTRAL EMITTANCES OF COMPLETE RADIATORS

Wien extended Boltzmann's method and showed that the wave-length for which the spectral emittance is a maximum is inversely proportional to the absolute temperature of the complete radiator. This law, which specified the displacement of the maximum of the spectral emittance curve as the temperature is varied, is known as *Wien's displacement law*. When λ_m is expressed in millimicrons and the absolute temperature in degrees Kelvin, the law[521] is given by:

$$\lambda_m T = 2.897 \times 10^6. \qquad (17)$$

Wien derived a corollary of his displacement law, that the maximum spectral emittance is pro-

[521] R. T. Birge, Rev. Mod. Phys. **13**, 233–239 (1941).

portional to the fifth power of the absolute temperature of a complete radiator:

$$W_m/T^5 = 1.29\times10^{-11}, \qquad (18)$$

in which W_m is expressed in ergs per second per square centimeter per millimicron.

Experimental spectral emittance curves for complete radiators at several temperatures are shown in Fig. 15.*,[522] As the temperature increases, the emittance increases rapidly and the wave-length at which the spectral emittance attains its maximum in each case shifts gradually into the region of shorter wave-lengths in accordance with Wien's displacement law. Wien extended his derivations to establish a general law which should hold for all spectral emittance curves for complete radiators:

$$W_\lambda = C\lambda^{-5}\cdot f(\lambda T). \qquad (19)$$

Wien attempted to derive the function $f(\lambda\cdot T)$ on the basis of an assumption concerning the detailed mechanism of the radiation of energy by oscillating particles. The formula resulting from this derivation, known as Wien's law,

$$W_\lambda = C_1\lambda^{-5}\exp(-C_2/\lambda T), \qquad (20)$$

predicts satisfactorily the spectral emittances for small values of (λT), but it deviates systematically from observed values when (λT) is increased. Because of its convenience for computation, Wien's law is still used to a considerable extent, with appropriate safeguards and corrections for the systematic errors.[523]

Rayleigh and Jeans derived a formula for the spectral emittance of complete radiators, on the basis of statistical mechanics and assumptions concerning the details of the radiation process which differed in several respects from those of Wien. The Rayleigh-Jeans law represents adequately the data for large values of (λT) which, however, are of less frequent occurrence in practice than the values of (λT) for which Wien's law is adequate.

The successful derivation of an accurate formula for the spectral emittance of complete radiators, by Max Planck in 1900, was one of the most spectacular and important accomplishments in the history of theoretical physics, and most of modern physical theory of atomic processes and radiation can be traced back to that event. Planck followed an argument that would lead to the Rayleigh-Jeans formula if carried to completion in the customary manner. To make the argument yield a formula which would fit the experimental data, Planck was forced to postulate that the energy ϵ of each of the elementary oscillators that constitute the complete radiator could not be subdivided indefinitely, but must be proportional to some integral multiple n of the frequency ν of the oscillator.

$$\epsilon = nh\nu. \qquad (21)$$

The most probable value[521] of *Planck's constant*

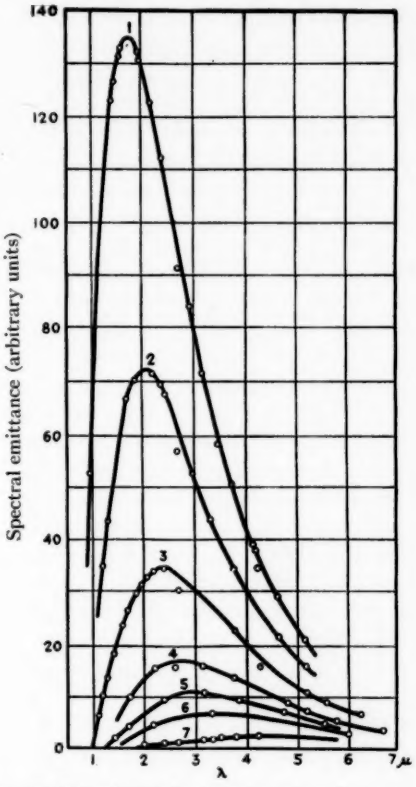


FIG. 15. Experimental emittance curves for a complete radiator (reference 522).
1—1646°K 3—1259°K 5—998°K 7—723°K
2—1449 4—1095 6—904

* Reproduced from *Scientific Basis of Illuminating Engineering*, P. Moon, by permission of McGraw-Hill Book Company, New York.

[522] O. Lummer and E. Pringsheim, Verh. d. D. Phys. Ges. **1**, 23–41, 215–235 (1899); **2**, 163–180 (1900).

[523] W. E. Forsythe, *Measurement of Radiant Energy* (McGraw-Hill Book Company, New York, 1937).

of action h is 6.624×10^{-27} erg second. Planck's formula for the spectral emittance is

$$W_\lambda = c_1\lambda^{-5}[\exp(c_2/\lambda T)-1]^{-1}. \qquad (22)$$

The most probable values [521] of the constants indicated in this formula are

$$c_1 = 3.740\times10^{9},$$
$$c_2 = 1.4385\times10^{7},$$

when W_λ is expressed in ergs per second per square centimeter per millimicron, λ is in millimicrons, and T is in degrees Kelvin.

Planck's formula fits the observed data within the limits of experimental error; it integrates into the Stefan-Boltzmann law for the total radiant emittance; it differentiates into the Wien displacement law; for small values of (λT) it becomes equivalent to Wien's formula; and for large values of (λT) it becomes equivalent to the Rayleigh-Jeans formula. Planck's law specifies the spectral emittance of complete radiators so completely and accurately that complete radiators are often called *Planckian radiators*. This synonym will be used freely in the following pages.

But, however adequately Planck's formula represents the facts, computations by the use of this formula are very laborious, and are usually avoided by use of tables[524–527] and charts[525] specifying relative values of spectral emittance for various temperatures and wave-lengths.* A method recommended by Forsythe[523] is to use the Wien formula, which is much more convenient for computation and fairly accurate for ordinary values of (λT), followed by the applications of corrections indicated by factors which Forsythe tabulated for various values of (λT).

[524] *International Critical Tables* (McGraw-Hill Book Company, New York, 1929) (c_2=1.433 cm deg., Vol. 5, pp. 238–242), S.A. p. 271.

[525] M. K. Frehafer and C. L. Snow, "Tables and graphs for facilitating the computation of spectral distribution by Planck's formula," Bur. Stand. Misc. Pub. 56 (1925) (c_2=1.435 cm deg.).

[526] J. F. Skogland, "Tables of spectral energy distribution and luminosity for use in computing light transmissions and relative brightnesses from spectrophotometric data," Bur. Stand. Misc. Pub. 86 (1929) (c_2=1.433 cm deg.).

[527] P. Moon, J. Math. and Phys. **16**, 133–157 (1937) (c_2=1.432 cm deg.).

* For precise work, the values given in these tables and charts should be corrected in accordance with changes in the most probable values (see reference 521) of c_1 and c_2 which have been recognized since the tables were computed.

19. EMISSIVITY

The complete radiator is unique in that its spectral emittance depends solely upon its temperature and not upon any of its other properties. Since the laws governing its behavior are so well known and have such a firm theoretical basis, it has become a sort of standard with which other less efficient radiators are compared. The *total emissivity* of an incomplete radiator is defined as the ratio of its radiant emittance to that of a complete radiator at the same temperature. If W represents the radiant emittance of the incomplete radiator, and if W_0 represents that of a complete radiator at the same temperature, the total emissivity e of the incomplete radiator is

$$e = W/W_0. \qquad (23)$$

When this equation is compared with Kirchhoff's law [Eq. (15a)] it is evident that the emissivity of any object is numerically equal to its absorptance. Since the absorptances of most substances depend upon both the wave-length of the incident light and the temperature of the absorbing surface, it follows that emissivity varies with both wave-length and temperature and also that emissivities for high temperatures in general cannot be determined from absorptances measured at moderate or low temperatures.

When the spectral emissivity of a radiator is constant throughout the spectrum, its spectral emittance bears a constant ratio to that of a complete radiator at the same temperature. Since the reflectance of such an object is a constant greater than zero, such an incomplete radiator is frequently called a "graybody." This name is confusing and objectionable, for the same reasons as the name "blackbody," and such radiators are called *non-selective radiators* in this report. Carbon is a fairly good example of a non-selective radiator, its spectral emissivity varying only from 0.8 to 0.5 in the wide spectral region between 0.4 and 4 microns.

A common example of a markedly selective thermal radiator is the Welsbach mantle, for which the emissivity varies from 86 percent at 450 millimicrons to 6 percent at 700 millimicrons. Incandescent filaments, however, have fairly uniform spectral emissivities in the visible portion of the spectrum. At 2400°K the variation

is from 46.3 percent at 467 millimicrons to 42.7 percent at 665 millimicrons. In the infra-red region of the spectrum, however, the emissivity of tungsten, like that of many other metals, drops rapidly to a low value. This selectivity, like that of the Welsbach mantle, is advantageous in its effect upon the luminous efficiency of such sources. It must be emphasized, however, that when the radiation is the result of thermal excitation alone, the emissivity can never exceed unity at any wave-length, since a complete radiator emits the maximum radiant flux possible at each and every wave-length.

20. RADIATION AND BRIGHTNESS TEMPERATURES

The concept of the complete radiator furnishes a convenient terminology for describing certain characteristics of radiating bodies. The *true temperature* of a body is, of course, its temperature on the Kelvin thermodynamic scale, read by a suitable thermometer in thermal equilibrium with the radiator. In the case of a complete radiator or a truly non-selective radiator the true temperature is the value of T for which Planck's formula accurately represents the spectral emittance. The *radiation temperature* of a body is the temperature of a complete radiator having the same total radiant emittance. The *brightness temperature*, or *blackbody temperature*, is the temperature of a complete radiator whose luminance ("brightness" in the older terminology) is the same as that of the body for some narrow spectral region, usually at 665 mμ. Strictly speaking it is a psychophysical concept, but since the band of wave-lengths transmitted by the filter used, and effective for vision, is fairly narrow, the concept can also be considered nearly equivalent to an abridged spectroradiometric specification. In terms of this concept, the *brightness temperature or blackbody temperature* is the temperature of a complete radiator whose radiant emittance (for some narrow spectral region, usually at 665 mμ) is the same as that of the object. Neither the radiation nor the brightness temperature can be greater than the true temperature. These "temperatures" are usually lower than the true temperature since spectral emissivities of all substances are less than unity.

21. COLOR TEMPERATURE

Most common radiators are fairly non-selective in the visible portion of the spectrum, and therefore the light from such illuminants can be approximately matched in chromaticity by the light from a complete radiator at some temperature. This temperature is known as the *color temperature* of the incomplete radiator, and the brevity and suggestiveness of this term has gained wide usage for it despite the considerable errors encountered in its determination and the frequency with which the concept is abused. It is often applied to selective radiators, such as Welsbach mantles and fluorescent lamps, and to radiators for which the term "temperature" has no meaning, such as the sky. Color temperature is a psychophysical concept and as such will be discussed in greater detail in later chapters. It

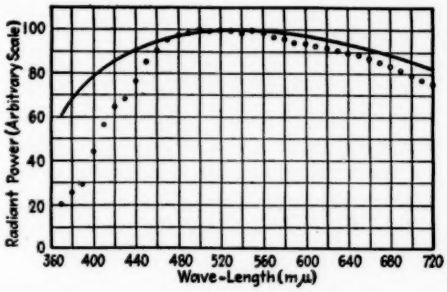


FIG. 16. The spectral distribution of mean noon sunlight at Washington in the visible and near ultraviolet regions (circles) compared with the emittance curve of a complete radiator at 5400°K.

is inappropriate for sources whose colors cannot be matched almost exactly by a complete radiator at some temperature. Even when a perfect color match is possible, a source cannot be assumed to have the same spectral distribution as the complete radiator with which the match is obtained. Sunlight, for example, corresponds to a color temperature of approximately 5400°K, but the spectral distribution of sunlight, Fig. 16,* does not resemble closely the spectral distribution of the energy from any complete radiator, even in the visible portion of the spectrum. Color temperature is a specification

* Reproduced from *Principles of Optics*, A. C. Hardy and F. H. Perrin, by permission of McGraw-Hill Book Company, New York.

of chromaticity, applicable to only a limited class of chromaticities, and is not a specification of spectral distribution. Grave errors and important discrepancies almost invariably result from the indiscriminate use of color temperature as a specification of the spectral distribution of a source. The color temperature of a non-selective radiator is the same as its true temperature. The color temperatures of other thermal radiators may be either higher or lower, depending on the variations of the spectral emissivities. The color temperatures of tungsten and many other metals are higher than their true temperatures. The light from a Welsbach mantle has a chromaticity so different from that of any complete radiator that color temperature is hardly an appropriate specification. When such specification is given, it might be called a *pseudo-color temperature*. The pseudo-color temperature of a Welsbach mantle is very much higher than its true temperature.

22. RADIOMETRY

When values of total emittance or of radiant transmittance or reflectance for spectrally heterogeneous energy are to be measured, we have seen that laborious computations can be avoided by using a non-selective detector and measuring total irradiance. Now detectors that convert radiant energy into electrical and chemical energy are highly selective, but certain materials absorb radiant energy almost non-selectively and thereby become heated. The problem is then simply to convert this thermal energy into a form that can be measured. Since thermal methods of radiation measurement *must* be employed in the infra-red, where electrical and chemical methods fail, the infra-red is sometimes regarded as the exclusive domain of thermal methods, but such methods must be used in *any* spectral region when absolute values are required. Inasmuch as absolute values are rarely of interest in colorimetry, we shall touch only lightly on thermal methods.

The thermocouple, which converts thermal into electrical energy, has attained a high state of perfection for radiometry. For this purpose, the thermocouple usually consists of a junction of bismuth and silver wire. A number of such couples can be connected in series and so disposed in space as to match the shape of the beam under investigation. The sensitivity increases only as the square root of the number of couples, however, which imposes a low practical limit. The thermopile is convenient because of its small size, but care must be taken to shield it from air currents and to avoid parasitic electromotive forces from other sources. The cell containing the pile is usually evacuated.

The bolometer consists of two fine platinum ribbons forming the arms of a Wheatstone bridge. One ribbon intercepts the flux, which upsets the balance of the bridge by reason of the increase in resistance. Like the thermopile, the bolometer must be protected from sources of parasitic electromotive forces, for which reason it is usually mounted in an evacuated chamber. It shares with the thermopile the advantages of stability and lack of susceptibility to mechanical disturbances but, because of its bulkiness, it is more sensitive to magnetic disturbances and is less convenient for some applications.

No radiometer can be better than the blackening of its receiving element. Soot has long been the standard material for coating, but Pfund has shown that its reflectance rises to as much as 50 percent at 11μ. He recommends the use of evaporated zinc or platinum, neither of which reflects or transmits as much as 1 percent when suitably deposited.

Apparently there is no tremendous difference in sensitivity between the various types of radiometer for construction of comparable quality. Cartwright[528, 529] has pointed out that the limit of the thermopile is 10^{-12} or 10^{-13} watt per square millimeter and that the limit of the others is of the same order of magnitude.

23. SELECTIVE DETECTORS

Over limited spectral regions photoelectric and photochemical detectors are used more commonly than thermal detectors because of their higher sensitivity. The greatest common disadvantage of such detectors is their marked selectivity, which prevents their use for absolute measurements except under special conditions, and which is likely to cause errors, even in

[528] C. H. Cartwright, Physics **1**, 211–229 (1931).
[529] C. H. Cartwright, Zeits. f. Physik **92**, 153–171 (1934).

comparison measurements, unless suitable precautions are used.

The most important selective detector, of course, is the eye. Inasmuch as subsequent articles are to be devoted to the phenomena of vision, nothing need be said here except that an observer cannot by direct observation, without the aid of instruments, determine the relative luminance of two fields which do not match. The eye is used very successfully in photometry to judge the equality of luminance of juxtaposed fields. The actual comparison is of the subjective attribute, brightness, and not of luminance. Under optimum conditions, however, such determinations are accurate with errors of less than one percent.

Detectors of radiant energy depending on photoelectric phenomena can be divided into three classes: photoconductive, photoemissive, and photovoltaic. The photoconductive class is represented almost exclusively by the selenium cell.[530] Such a cell requires an external source of potential and it functions by virtue of the change of its resistance when irradiated. The resistance of a selenium cell is low, so its output cannot easily be amplified, and its response is proportional only to the square root of its irradiance. It responds only to energy in the near infra-red region of the spectrum.

Photoemissive detectors consist of surfaces of alkali metals or their compounds which emit electrons when they are irradiated. The photoemissive surfaces are enclosed in an evacuated glass (or quartz) bulb, together with a collecting anode. When a potential difference of approximately 100 volts is applied between the photoemissive cathode and the anode, the current produced is practically proportional to the incident flux, the deviations from linearity being small even for high values of flux and potential difference. The sensitivity for low values of irradiance can be greatly increased by the presence of a rather low pressure of an inert gas, which is ionized by collision with the emitted electrons, but the departure from linearity is then marked for high values of flux and potential. Moreover, such a cell can be seriously damaged by a glow discharge which may occur at high values of irradiance and potential.

An important law regarding the spectral sensitivity of a photoemissive cell can be deduced by noting that the kinetic energy of a photoelectron is proportional to the frequency of the radiant energy that excited it. A certain amount of energy (called the work function) is required for the electron to escape from the emitting surface. Unless the electron receives this amount of energy from the incident energy, it cannot escape, which indicates that a photoemissive cell has a low frequency (long wave-length) limit of response. Thus sodium, which has a large work function, is used for the cathodes of cells designed to respond to the ultraviolet and not to the visible. To lower the work function, the alkali metals are compounded with sulfur, oxygen, or hydrogen. Thus the potassium-oxide cell is excellent for the visible region, while the caesium-oxide cell, which has supplanted all others for ordinary applications because of its sensitivity, is responsive to radiant energy with wave-lengths as long as one micron. It has one sharp maximum of sensitivity at 350 mμ and a broad maximum, almost as great, at 800 mμ.

Photoemissive currents are of the order of a few microamperes, but the response can be increased readily by vacuum-tube amplifiers because of the high resistance of such cells. The effect of stray radiant energy can sometimes be reduced by choosing a cell of suitable spectral sensitivity. When this is not sufficient, the beam under investigation can be interrupted with a chopper and the current amplified and filtered with an alternating-current amplifier. Under ordinary conditions of use, photoemissive cells exhibit little lag (vacuum cells none at all), nor are they subject to fatigue save on exposure to a powerful beam of radiant energy.

A photoemissive cell can be placed anywhere in the beam, provided that excessive local heating is avoided because its response is proportional to incident flux, almost regardless of the spatial distribution of the flux. Photoemissive cells retain their calibration fairly well, but frequent checks are advisable. The use of a null method, in which the comparison and the unknown beams are rapidly interchanged and the comparison beam varied until the difference

[530] G. P. Barnard, *The Selenium Cell, Its Properties and Applications* (Constable, London, 1930).

between the responses is reduced to zero, is preferable to methods which rely on the calibration of the cell. The sensitivities of photo-cells may vary slightly but appreciably from one portion of the cathode to another, so it is advisable to have the same portion of the cathode irradiated by both the comparison and the sample beam. The sensitivity of a typical caesium-oxide vacuum cell to flux from a tungsten source at a color temperature of 2870°K is about twenty microamperes per lumen, while the corresponding response for the gas filled cell is about one hundred microamperes.

The electron multiplier, which is now coming into considerable use, consists essentially of a photo-cell and an amplifier in a single envelope. The amplifier works on the principle that the photoelectrons cause the emission of secondary

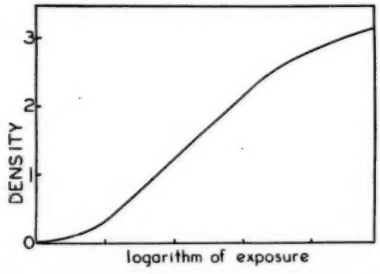


FIG. 17. Characteristic curve of photographic material.

electrons by collision. The output is measured in terms of milliamperes, but a magnetic or electrostatic field of great constancy is required, which calls for elaborate auxiliary apparatus.

The photovoltaic, barrier-layer, or photo-e.m.f. cell, as it is variously termed, consists of a metallic plate covered with a thin layer of a semiconductor, such as cuprous oxide, on which is deposited a film of metal. This film is thick enough to serve as an electrode but thin enough to transmit considerable radiant energy, which generates a potential of a few millivolts in the barrier layer. The resistance of the layer decreases simultaneously, so the current increases linearly with incident flux provided the external resistance be of the order of only a few ohms. This type of cell is popular for portable instruments because it requires no external source of electromotive force and its output can be measured with a microammeter, the sensitivity of a representative cell being 145 microamperes per lumen. Its dark current is zero, it has no sensible lag, and its calibration is satisfactorily stable for ordinary purposes despite powerful irradiation. High temperature is deleterious, however, the safe limit being about 50°C (120°F). Furthermore, the sensitivity of a photovoltaic cell may not be uniform over its surface. The spectral sensitivity of a typical cell has a maximum at 580 mμ but extends far into the ultraviolet and infra-red. A combination glass filter is available that produces a close match to the luminosity curve of the eye,[531] when used with some of these cells. This filter is not appropriate for other cells which have markedly different spectral sensitivities.

The photographic plate is almost the only type of photochemical detector in practical use. Some prejudice concerning its use in quantitative work has arisen as a result of attempts to employ it without proper consideration for its characteristics, but it is convenient, rapid, and capable of measurements of high accuracy when appropriate precautions and methods are used.[532, 533] The sensitivity of a plate is an inverse measure of exposure (time of exposure multiplied by irradiance or illuminance) required to produce a certain response, and depends upon such things as the temperature and moisture content of a plate during storage and exposure. It is even affected by the length of time elapsing between exposure and development, the latent image growing or decaying according to the type of plate in an unpredictable manner.

To an approximation satisfactory for ordinary photographic purposes, the density produced varies as the product of irradiance and time regardless of the relative magnitudes of these factors. This product is defined as *exposure* and a typical density *versus* exposure curve is shown in Fig. 17. Departures from this reciprocal relationship of irradiance and time for constant density are too great to permit its use in photometric measurements, especially for extremely high or extremely low values of irradiance.

[531] H. P. Gage, J. Opt. Soc. Am. **27**, 159–164 (1937).
[532] G. R. Harrison, J. Opt. Soc. Am. **19**, 267–316 (1929).
[533] L. A. Jones and J. W. McFarlane, J. Soc. Mot. Pic. Eng. **19**, 361–390 (1932).

Similarly, the density produced by an intermittent exposure with a constant value of irradiance during the exposure cycles is intermediate between the densities produced by the same total energy received in uninterrupted exposures at the instantaneous and at the average rate. The density produced by the intermittent exposure is equal to the density produced in the same total time by the same average irradiance only when the frequency of the intermittent exposure exceeds a critical frequency corresponding to less than one quantum per flash per grain in the photographic material.[534] Provided that this critical frequency is exceeded, a rapidly rotating sector disk can be used to control the effective irradiance in photographic photometry.

All photographic materials are sensitive from about 500 mμ to 250 mμ, where the absorption in the gelatin becomes serious. For shorter wave-lengths, most of the gelatin must be eliminated, or a layer of a fluorescent substance, such as some oils or minute crystals of certain organic compounds, must be applied to the photographic plate. For longer wave-lengths, certain dyes are incorporated into the photographic emulsion, and plates sensitive to 1.2 microns are now available.[535]

Very marked and erratic variations in density can arise from variations in the processing technique, of course, and although this technique can be refined so as to secure an amazing degree of reproducibility, comparison exposures should be included on each plate for the most precise photographic photometry. Moreover, since there may exist minute differences in sensitivity over a single plate, the comparison and test exposures should be made on adjacent small areas, and they should be repeated in the same manner several times.[532, 533] In other words, whenever practicable, photographic materials should be used in photometry merely for detecting and recording equality or lack of equality of beams of identical spectral composition, in a manner equivalent to that in which the eye is employed in photometry. Photographic methods are far superior to the eye or to any other type of detector for extremely low values of irradiance because the irradiance can be integrated photographically over long periods of time. The precision of photographic photometry, properly conducted, is limited only by the precision with which the densities of contiguous areas can be compared.

The type of detector to be employed is of paramount importance in the design of any optical instrument. Efficient results cannot be obtained by merely taking a visual instrument and introducing a photographic plate, or a thermopile, or a photo-cell at the eye point. A tabulation of the characteristics of radiant energy that are significant in connection with the various types of detectors will make this clear.

Detector	*Response depends upon:*
Human eye	Spectral radiance or luminance
Photoconductive cell	Spectral irradiance
Photoemissive cell	Spectral flux
Photovoltaic cell	Spectral flux
Photographic plate	Spectral energy density
Non-selective detectors	Radiant flux.

Since the response of any selective detector depends upon the spectral distribution of the energy incident upon it, a standard distribution is usually stated (or assumed) in specifications of over-all sensitivities. For instance, in the specification of the sensitivisties of photographic materials, and in the calibration of photoelectric exposure meters for photography, a standard artificial daylight source is employed.[536, 537] Anomalous results are occasionally obtained, therefore, when exposure meters and photographic materials are used with sources such as some phases of natural daylight which differ considerably from the spectral distribution of the standard calibration source. As in the case of most selective detectors, the spectral sensitivities of the photographic materials and of the photocell in the exposure meter are rarely known. Consequently, a quantitative discussion of the effects of differences between the spectral distributions of various sources would not be useful. For the purposes of general, qualitative discussions, however, a concept of radiant purity is convenient. Radiant purity can be defined in a

[534] J. H. Webb, J. Opt. Soc. Am. **23**, 157–169 (1933).
[535] *Photographic Plates for Spectroscopy and Astronomy* (Eastman Kodak Company, Rochester, New York).

[536] L. A. Jones, J. Opt. Soc. Am. **21**, 361–368 (1931).
[537] Report of the O. S. A. Committee on the photographic standard of intensity, J. Opt. Soc. Am. **21**, 654–676 (1931).

manner that is suggestive and simple, although highly arbitrary. The quantity thus defined should be considered as no more than a crude classification index, related to the probability that any particular spectral distribution will result in anomalous relative sensitivities of various selective detectors.

The radiant purity of a distribution represented by the curve E_λ (with maximum value E_m) in any wave-length region of interest (between λ_1 and λ_2) is the area between the curve E_λ and the horizontal line at E_m, from λ_1 to λ_2, divided by the total area under the line E_m, from λ_1 to λ_2. In Fig. 18 the radiant purity is the ratio of the total shaded area, marked A, to the area of the rectangle under E_m between λ_1 and λ_2. Analytically, radiant purity S of the distribution E_λ (maximum E_m), from λ_1 to λ_2 is defined by the integral:

$$S = \int_{\lambda_1}^{\lambda_2} (1 - E_\lambda/E_m) d\lambda.$$

Radiant purity as thus defined has no fundamental physical basis. The concept is purely arbitrary and is useful only because it is suggestive and convenient for qualitative discussions of the behaviors of selective detectors. Like any other crude classification index, the quantity defined as radiant purity is an incomplete representation of facts which can be represented completely only in terms of a large number of data, such as are represented by the curve E_λ, in Fig. 18. Anomalous ratios of the sensitivities of selective detectors are more probable for high radiant purities than for low.

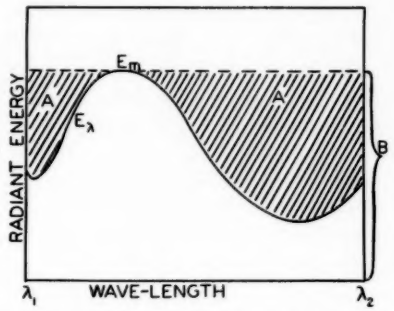


FIG. 18. Definition of radiant purity: ratio of area A to area of rectangle under E_m between λ_1 and λ_2.

24. THE MODULATION OF RADIANT ENERGY

One of the most difficult problems in the handling of radiant energy is to control its flux density in a known and reproducible manner when a high degree of precision is required. Varying the input to the source is rarely feasible because of the alteration of the spectral composition of the output. Perhaps the most reliable method is to vary the distance of the source and compute the irradiance on the basis of the inverse-square law, which was described by Bouguer in 1729.[538] Even this is unreliable if the dimensions of the source are not negligible compared with the distance from the source to the plane on which the irradiance is measured or if appreciable absorption occurs in the intervening medium.

When the source has a considerable area, like a sheet of opal glass, its output can be controlled by varying its area, but the radiance must be uniform and the computation of irradiance as a function of area of the source is complicated when its dimensions are appreciable fractions of the distance to the irradiated surface. In an image-forming system, a diaphragm can be placed at the aperture stop. The familiar iris is very convenient and provides a continuous variation, but, unless the zonal errors of the lenses are insignificant, the irradiance at the image plane will not be strictly proportional to the area of the iris. To avoid the effect of zonal aberrations, the diaphragm can be made of two sector disks mounted on a common axis so that one can be rotated with respect to the other, the variation being between zero and one-half of the area of the aperture with the diaphragm removed.

In some spectrophotometers luminance of the comparison field is controlled by varying the width of the entrance slit, the eye being focused on the telescope objective. Proportionality between slit width and luminance should not be assumed without verification. Lens aberrations and diffraction effects cause discrepancies in this method.

Filters are convenient and are of general applicability, but truly non-selective ones are difficult to obtain. "Neutral" glass and dyed gelatin filters are far from neutral even in the

[538] M. Bouguer, *Essai d'optique sur la gradation de la lumière* (Jombert, Paris, 1729).

visible and their transmittance is usually much higher in the infra-red than in the visible. Recently developed gelatin filters containing colloidal carbon are superior to the older dye neutrals, but carbon suspensions increase in transparency in the infra-red beyond about 1.2μ and, furthermore, the carbon particles are large enough to impair slightly the definition of an image. Platinum sputtered on glass has been recommended as being non-selective in the visible and ultraviolet and should also be suitable for the infra-red, but the character of the deposit depends upon the conditions of deposition. A filter consisting of wire screening is non-selective and may be used in place of iris or sector diaphragms, but must be calibrated in the system in which it is to be used.

A sector disk rotating above the critical frequency has been found to reduce the luminance of a photometric field exactly in proportion to the ratio of the open sectors to the entire surface of the disk.[539] The photographic plate also integrates intermittent exposure, provided its critical frequency is exceeded. This frequency should be determined for the type of plate employed by making a series of tests at various frequencies. The chief disadvantage of the sector disk is that it cannot be varied while in motion except by use of elaborate mechanisms that are likely to be noisy and distracting. A disk whose angular opening varies from the center to the periphery, like that designed by Hyde,[540] is an elegant device which is useful when the beam has a very small cross-sectional area. The transmittance of such a disk can be varied while it is rotating by moving it radially across the beam.

When it is necessary to vary the irradiance from one side to the other of an area, as in a wedge spectrograph,[541–543] wedge-shaped filters can be used under the limitations of neutrality mentioned in connection with uniform filters. Wedge or step diaphragms can be placed at the aperture stop and imaged on the slit of a spectrograph with a cylindrical lens.[544] This method, however, is not suitable for precise work because of the aberrations of the lenses, but it is adequate for wedge spectrographs, which are employed primarily for exploratory work, and for such a purpose precision can be sacrificed to simplicity and convenience.

25. SELECTIVE FILTERS

The spectral composition of radiant energy can be modified either by means of substances that absorb selectively within the spectral region under consideration,[545] or by dispersing the beam and recombining the portions of the desired wavelengths. Only the first method will be considered in this section, the second being treated in connection with other dispersing instruments.

Common colored glass cannot ordinarily be used, even when there is no question of ruining an image with striae, because the spectral transmittance of such glass is rarely suitable. Series of colored and of fairly neutral glasses are made especially for optical use by both Schott[546] and Corning.[547] These colored glasses have cut-offs that approach the sharpness of dyes. A filter can be made conveniently with the simplest equipment by fixing out an undeveloped photographic plate and soaking it in a suitable dye. This method was once popular, but it is practically superseded by the Wratten filters,[548] which consist of sheets of dyed gelatin. They can be obtained in the form of unmounted film, or cemented between glass. Gelatin filters are readily damaged by heat and moisture. Glass filters are less likely to suffer permanent modification by heat and are not ordinariily damaged by moisture. Some filters fluoresce, but this phenomenon is rarely so pronounced as to be objectionable.

In general, transmittance bands are sharper on their short wave boundaries than at the long wave limits. Consequently, most filters that absorb only the short wave portions of the visible spectrum have a sharper cut-off than those which

[539] E. P. Hyde, Bull. Bur. Stand. **2**, 1–32 (1906).
[540] E. P. Hyde, Astrophys. J. **35**, 237–267 (1912).
[541] C. E. K. Mees and S. H. Wratten, Brit. J. Phot. **54**, 384–385 (1907).
[542] T. R. Merton and J. W. Nicholson, Phil. Trans. **A217**, 237 (1917).
[543] J. K. Robertson, J. Opt. Soc. Am. **7**, 983–995 (1923).
[544] O. E. Miller, Rev. Sci. Inst. **3**, 30–42 (1932).
[545] K. S. Gibson, J. Opt. Soc. Am. **13**, 267–280 (1926).
[546] Schott and Gen., *Jena Colored Optical Filter Glasses for Scientific and Technical Purposes*, Jena Glass Works (U. S. agents: Fish-Schurman Corporation, 230 East 45 Street, New York).
[547] Corning Glass Works, *Glass Color Filters* (Corning, New York).
[548] Eastman Kodak Company, *Wratten Light Filters* (Rochester, New York).

absorb primarily the long wave portions of the visible spectrum. Similarly, it is much easier to absorb the ultraviolet than the infra-red, without seriously affecting the visible energy. Several types of filter, of which Noviol is a well-known example, are available for the removal of the ultraviolet. The converse problem, of transmitting the ultraviolet, is much more difficult. Quartz and fused silica are commonly used, but certain glasses whose transmittances are high in the near ultraviolet are now available in large sheets. Corex glass is exceptional, its transmission band extending to 220 mμ, and a variety that is practically opaque in the visible is made. The short wave limit of transmission of ordinary glass is about 300 mμ for two-millimeter thickness; in thicknesses such as are used in most optical instruments this limit is nearer 350 mμ.

Although glass and gelatin filters generally absorb the ultraviolet, most of them transmit the infra-red. Certain glasses, such as Aklo, absorb the infra-red while transmitting the visible spectrum to a high degree, but they are likely to crack when strongly irradiated unless heat-resistant. Although frequently inconvenient, water is a readily available absorber for the infra-red, a two-centimeter layer transmitting little beyond 1.2 microns. Solutions of copper salts are more effective, but their absorption bands extend into the visible region, the solutions appearing greenish. Coblentz has recommended a 2.5 percent solution of cupric chloride. A thin layer of gold deposited on glass transmits 80 percent of the luminous flux from an incandescent lamp, but only 20 percent of the infra-red energy.

A problem of much importance in visual photometry is the modification of the output of one source to match that of another in chromaticity. Two series of Wratten filters[548] are made for this purpose, one to raise the color temperature and the other to lower it. Two-part liquid filters containing copper sulfate and cobalt ammonium sulfate have been designed by Davis and Gibson[549] to raise the color temperature of light from incandescent sources operating at various temperatures. Certain of these have been adopted internationally for producing standard artificial daylight for colorimetry and photographic sensitometry.[536, 537] Corning Daylite filters (glass) are suitable for large permanent artificial daylighting installations.[550] When the output of a source is to be modified to match a certain color temperature, it must be well considered whether a mere visual match is adequate or whether the spectral distribution must be matched.

Both glass and gelatin filters are also available for isolating narrow spectral regions, but their maximum transmittance is so low that line sources used with less selective filters that isolate a single line are frequently preferable. The Christiansen filter,[523, 551–553] which consists of a powdered transparent solid immersed in a liquid of similar index but widely different dispersion, is an interesting device for isolating narrow spectral regions at will. In the region where the index of solid and liquid is the same, the beam is transmitted with but slight loss; in other regions, the indices differ and most of the energy is scattered outside the beam. By using a mixture of liquids, the location of the transmission band can be controlled. The difficulty with this type of filter is to get a solid of sufficient homogeneity. Natural crystals are uniform but expensive. Ordinary glass is entirely useless, and even optical glass is not satisfactory unless it all comes from a single small block.

26. PHOTOMETRY

Radiant energy in the visible portions of the spectrum has usually been evaluated in terms of the luminous sensation it produces when it falls upon the eye of a human observer. The observer is required to judge equality of brightness. Experience has proven that evaluations of luminance by equating brightness are sufficiently precise for accurate measurements. Methods employing the eye in this manner are rapid, require relatively simple and inexpensive apparatus, and yield the most significant results for visual applications. The economic importance of photometry

[548] R. Davis and K. S. Gibson, "Filters for the reproduction of sunlight and daylight and the determination of color temperature," Bur. Stand. Misc. Pub. 114 (1931).

[550] H. P. Gage, Trans. I. E. S. **31**, 995–1022 (1936).

[551] C. Christiansen, Ann. Phys. u. Chem. **23**, 298–306 (1884).

[552] E. D. McAlister, "The Christiansen light filter; its advantages and limitations," Smithsonian Misc. Pub. 3297 (1935).

[553] R. B. Barnes and L. G. Bonner, Phys. Rev. **49**, 732–740 (1936).

is so great that rigid standards have necessarily been in use for many years, and it is not the purpose of this report to discuss such standardization[13,*] except where the information is of direct use in colorimetry. This report would not be complete, however, without a brief description of photometric methods. Although the quantities measured are psychophysical, the instrumental procedures are physical[554] and so it is convenient to discuss the practical methods of photometry in this section.

The first photometric standard adopted was the standard of intensity, which consisted of a candle made and burned according to definite specifications. The procedure involved in comparing a known source, such as the standard candle, with an unknown is to place the sources

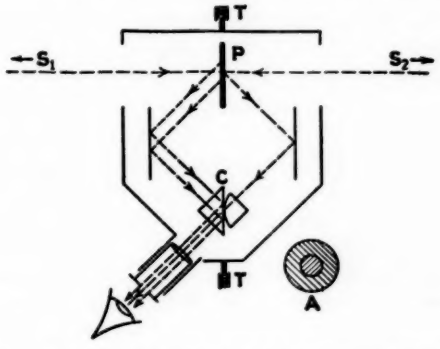


FIG. 19. Sketch showing the principle of Lummer-Brodhun photometer head.

at the ends of a track and then determine the intermediate point where they produce the same illuminance. In most types of photometer, each source exclusively illuminates a white diffusing surface, and the surfaces thus illuminated are viewed in juxtaposition. To mention an example that is still used because of its simplicity, the Ritchie wedge consists of two white diffusing surfaces which intersect at an angle of ninety degrees, the bisector of this angle being normal to the line joining the two sources. For highest precision, the edge must vanish at the balance point, and in this respect the Lummer-Brodhun cube surpasses the Ritchie wedge. This cube consists of two prisms, as shown in Fig. 19,* clamped together in optical contact. Light from the screen S_1, illuminated by one source, passes directly through the center of the prisms where they are in contact, while that from the screen S_2, illuminated by the other source, is reflected into the ocular from the portions of prism P_2 where the prisms are not in contact. The remaining portions of the light from the two screens are directed away from the ocular and are absorbed.

The inverse-square law, which is used on all bench photometers, is accurate within three-tenths of one percent provided that the greatest dimension of either source is less than eleven percent of the distance from that source to the photometer screen.

Increasing demands for higher accuracy caused the abandonment of the standard candle in favor of flame standards of various kinds, the last of which (an amyl-acetate standard devised by von Hefner-Alteneck in 1884) has only recently been abandoned in Germany. The flame standards were almost universally replaced by a group of incandescent lamps used by the national standardizing laboratories to calibrate working standards. The use of a group of lamps as a standard had the disadvantage that the standard could not be reproduced from specifications. Consequently, an international agreement to use one sixtieth of the luminance of a complete radiator at the freezing point of platinum (approximately 2044°K) as the fundamental photometric standard was adopted,[555] to go into effect† on January 1, 1940. (On the basis of the former international candle based on groups of lamps, the luminance of such a complete radiator is 58.9 lumens per square centimeter.[556])

The physical standard consists of the interior of a heat-resistant (thoria) tube immersed in a

* Reference numbers from 1 to 16 indicate publications listed in the first article of this series, entitled "The historical background and evolution of the colorimetry report," J. Opt. Soc. Am. **33**, 534–543 (1943).

[554] J. W. T. Walsh, *Photometry* (Constable, London, 1926).

* Reproduced from *Principles of Optics*, A. C. Hardy and F. H. Perrin, by permission of McGraw-Hill Book Company, New York.

[555] E. C. Crittenden, J. Opt. Soc. Am. **29**, 103–107 (1939).

† The required exchange of working standards calibrated against this fundamental standard, by various national standardizing laboratories, was not completed before the outbreak of war in 1939. Consequently, the new photometric standard is not yet in use, and probably will not be put into use until the exchange of working standards can be completed satisfactorily (reference 13).

[556] H. T. Wensel, W. F. Roeser, L. E. Barbrow, and F. R. Caldwell, Bur. Stand. J. Research **6**, 1103–1117 (1931).

vessel of pure, molten platinum in an induction furnace.[556] The comparison is made by focusing an image of the tube on a photometric screen and determining the photometric balance point as the molten platinum solidifies. Suitable allowance is made for the light loss at the lens. It goes without saying that this procedure is rarely performed and then only to compare the standard with more convenient incandescent lamps that serve as secondary standards.

In the measurement of the luminance of the radiator just described, the opening in the tube is so small that an enlarged image of the interior wall has to be projected on a screen for observation. It is more usual, however, to deal with surfaces large enough so that they can be viewed directly through the Lummer-Brodhun cube in place of the screen used for measuring sources. The comparison surface is a diffusing screen illuminated by a light source whose intensity can be controlled. This control is usually performed by varying the distance of the source from the comparison surface or by moving a neutral wedge. A similar apparatus will serve to measure the illuminance by observing a diffusing plate of known reflectance placed at the plane where the value of illuminance is desired.

Modern lamps are used in such a way for ordinary lighting that the value of total flux is more significant than is the candlepower in a particular direction. Total flux can be determined by the laborious procedure of measuring the candlepower at various altitudes above and below the lamp's equator and then integrating, taking the relative areas of the various zones into account. By far the simpler method is to place the lamp in a large sphere with a white, diffusing interior wall and measure the luminance of the wall by means of a photometer set at a hole. It was shown by Sumpner that the luminance of the wall of such a sphere is proportional to the total flux emitted by the source within, and Ulbricht in 1900 proposed the use of such an integrating sphere for photometry. In practice, a source of known output is introduced and a reading made to calibrate the comparison source. More commonly in industrial laboratories, the output is determined by means of a photoelectric cell equipped with suitable filters.[523]

These common methods fail when the source to be measured is comparatively small or weak, but then the principle of the Maxwellian view can be applied. This consists of focusing the source on the eye, in which case the lens, which appears uniformly bright, serves as one of the photometric screens.

So far nothing has been said of the chromaticities of the two fields under comparison: When the chromaticities are different, the problem of equating luminances is, to quote an eminent French photometrician, "embarrassing." Visual difficulty may be minimized, however, by provision of an acceptable judgment situation. Sometimes the chromaticities can be matched merely by introducing one or more filters, in which case their presence must be allowed for by the computational methods which will be described in a subsequent article. Two direct methods are available when the chromatic differences are large. The step-by-step method bridges the interval with a series of sources intermediate in chromatcity. The flicker method presents the two beams to the eye in rapid alternation, the balance point being taken as the point where the flicker is a minimum. The virtue of the flicker photometer resides in the psychological fact that the critical frequency for chromatic flicker is lower than for luminance flicker; the minimum for luminance flicker can thus be determined without the annoying presence of chromatic flicker. It is important that the transition between the sources be made smoothly so that no extraneous flicker will be introduced.

The radiant flux emitted by a hot body can often be taken as an index of the temperature of the body.[523] One obvious method of measuring this flux—radiometric rather than photometric—is to focus a beam of known angular extent on a non-selective detector. The photometric method is to determine the brightness temperature at some spectral region (usually about 665 mμ). The usual instrument for doing this is the optical pyrometer which consists essentially of a telescope in the focal plane of which is the filament of a small incandescent lamp. The current through this lamp is adjusted until the filament disappears against the background of the hot body, whereupon the resistance of the filament or the current through it is taken as a measure of its brightness temperature by means of an em-

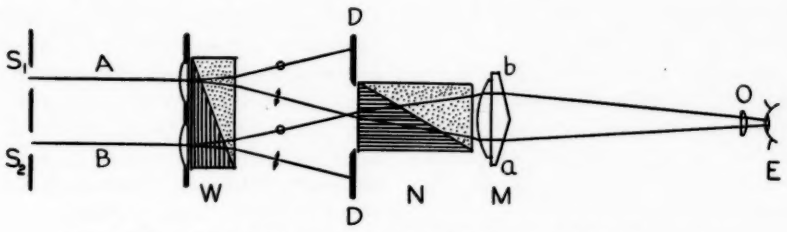


FIG. 20. Sketch showing the principle of the Martens polarizing photometer.

pirical calibration curve. The brightness temperature can alternatively be measured with a photo-cell. Unless the source under investigation is a complete radiator, as is approximately the case for the interior of a furnace, corrections must be applied to the readings of an optical pyrometer.

The methods of measuring reflectance and transmittance depend upon the conditions under which the determinations are required. When the reflectance of a surface in a given direction is desired, the luminance of the surface in that direction under the specified conditions of illumination can be compared with the luminance of a surface of known directional reflectance. When diffuse reflectance is required, the surface can be placed against an opening in an integrating sphere and illuminated with a beam that does not directly illuminate any other part of the sphere. The luminance of the sphere wall is then proportional to the reflectance of the sample, and the proportionality factor can be determined by measuring the luminance of the sphere wall when the incident beam is directed onto an adjacent portion of the sphere wall for which the diffuse reflectance is already known. Errors will be minimized if the sample is left in place in the opening of the sphere during this calibration of the sphere. Substitution of a reflectance standard in place of the sample for the calibration will result in errors proportional to the fraction of the sphere wall occupied by the sample and proportional to the difference of the diffuse reflectances of the sample and standard.[557] Diffuse transmittance can be measured in an analogous manner.

Specular transmittance can be measured by introducing the sample into one side of a photometer. The beam passing through the sample should be collimated so that all rays traverse the normal distance between the faces of the sample. When the inverse-square law is used for the photometric comparison, the difference between the physical thickness and the optical thickness of the sample may be appreciable for thick samples and allowance should be made for this effect. When the sample is a solution, it is customary to compare the solution against the solvent placed in an optically equivalent container.

The measurement of transmittance and reflectance is facilitated if the beams illuminating the sample and standard are taken from the same source. The Martens polarization photometer sketched in Fig. 20 is especially convenient for such measurements. In this instrument two beams, *A* and *B*, fall on adjacent areas of a Wollaston prism *W*. Each beam is split into two components, one of which is refracted so as to pass through an analyzing Nicol prism *N*. The other component is refracted away from the aperture of the Nicol and is removed by a diaphragm *D*. The components of the two beams which are incident on the Nicol prism are plane-polarized in perpendicular azimuths, and the fractions of these two components transmitted by the Nicol prism are equal to the squares of the sine and the cosine of the angle between the principal plane of the Nicol and the plane of polarization of one of the components. The portion of beam *A* which ultimately falls on face *a* of the biprism is refracted into the eye *E* of the observer, as is also the portion of the beam *B*, which ultimately falls on face *b* of the biprism. The portions of beam *A* which fall on face *b* of the biprism, and of beam *B* which fall on face *a* of the biprism, are refracted away from the eye, so that the light which reaches the eye of the

[557] A. C. Hardy and O. W. Pineo, J. Opt. Soc. Am. **21**, 502–506 (1931).

observer from each half of the biprism is derived from only one of the original beams. The refracting edge of the biprism is at the focus of the ocular lens *O* and forms a sharp dividing line between the two halves of the photometric field. The orientation of the analyzing Nicol, when the photometric field is balanced, indicates the relative luminances corresponding to the original beams falling on the Wollaston prism. This ratio is given by the square of the tangent of the angle between the principal plane of the Nicol and one of the planes of polarization of the components incident on the Nicol. If two beams derived from a single source, or from sources having equal luminances, are directed towards the Wollaston prism, and an absorbing object is placed in one of these beams, the transmittance of the sample will be given by the square of the tangent of the angle just defined, when the photometric field is balanced. The Martens polarizing photometer can be used for the measurement of reflectance by an appropriate arrangement of the sample and a standard of known reflectance.

The Martens type of photometer is adaptable for a great many purposes and is deservedly popular but, in addition to sources of errors which are encountered in all forms of photometers, polarizing photometers are susceptible to errors arising from unsuspected polarization of the samples, as discussed in Section 16.

Instruments designed especially for measuring the densities of photographic deposits are called densitometers. One of the most reliable visual types consists of a Martens photometer arranged so that one beam passes through the sample. Diffuse densities are usually required, in which case the sample is illuminated by placing an illuminated opal glass plate in contact with the sample, and the comparison beam is derived from an adjacent area of the same opal glass not covered by the sample. The opal glass should be sufficiently dense to produce adequate diffusion of the incident light. A slightly more perfect result could be obtained by illuminating the sample by placing it against an opening in an integrating sphere.[513]

A calibrated "neutral wedge" of which the density varies continuously from point to point is frequently used in densitometers, especially when many densitometers are to be constructed at low cost, or when the area for which the density is to be determined is so small that a practical polarization photometer cannot be designed. Accurately non-selective wedges are rare and difficult to manufacture, and most wedge photometers exhibit differences in the chromaticities of the two portions of the photometric field. These differences may become quite serious for some densities, and accurate photometric comparisons are difficult in the presence of such differences.

Accurate visual photometry is so difficult and tedious that methods employing physical detectors have been extensively investigated. One of the first methods employed a thermopile, with a filter whose spectral transmittance was approximately proportional to the luminosity curve for the human eye. But the low sensitivity of non-selective detectors has diverted attention to selective detectors which are more sensitive. The use of a photovoltaic cell with a correcting filter is mentioned in Section 23. Photoemissive cells are advantageous because the responses of such cells can be amplified conveniently, but the auxiliary apparatus required for such instruments is so much more complicated and usually less portable than that required by photovoltaic cells, that the latter are used more frequently at present. When considerable variations in spectral composition are not encountered, satisfactory approximations to the luminosity curve can be secured with filters designed specifically for each type of photovoltaic or photoemissive cell.[531] Photo-cells corrected to luminosity characteristics with filters are used extensively for routine photometry of incandescent lamps, and for transmittance and reflectance measurements. Photoelectric densitometers usually indicate transmittance on a linear scale, having very open scales of density at low densities and undesirably crowded scales for high densities. Logarithmic amplifiers can be employed to convert the photoelectric response (which is proportional to transmittance) to an electrical signal which is linearly related to density.[558] Substitution methods, in which calibrated densities are substituted for the sample until equal responses are obtained, or methods in which calibrated densities are super-

[558] J. P. Taylor, Electronics **10**, 24–27 (1937).

imposed over the sample until a standard photoelectric response is obtained,[559] are also used successfully for photoelectric densitometry.

27. DISPERSING SYSTEMS

No single name is in general use to cover the entire class of dispersing instruments. They all consist of a slit, a collimating lens (or mirror), a dispersing element, and a second lens (or mirror) to bring the spectrum to a focus at a convenient location. When the spectrum is examined with an ocular, the instrument is properly called a *spectroscope*, but this name is also frequently used to refer to the whole class. A spectroscope fitted with a divided circle and facilities for readily interchanging prisms, so that the index of refraction of the prism can be determined, is called a *spectrometer*. When the ocular lens of a spectroscope is replaced with a photographic plate, the instrument is called a *spectrograph*. When an exit slit is used in place of the ocular, for the isolation of narrow portions of the spectrum, the instrument is called a *monochromator*. When a monochromator is combined in any manner with a photometer, for the measurement of relative spectral emittances, spectral transmittances or reflectances, the instrument is called a *spectrophotometer*.

The basic type of prism is the simple 60-degree type sketched at *A* in Fig. 21. The deviation of the emergent ray from its direction when incident is at a minimum when it is normal to the bisector of the refracting angle as it traverses the prism. The Fresnel loss is then at a minimum also, and the aperture of the prism is at a maximum. This type of prism is suitable when the direction of the emergent ray may vary with wave-length, as when the spectrum is to be explored visually; it is not suitable when the deviation must be constant while the wave-length corresponding to this deviation is varied. This last requirement must be fulfilled in monochromators, the spectrum sweeping past the exit slit as the prism is rotated. The Pellin-Broca prism, sketched at *B* in Fig. 21, is suitable for this purpose, and is the type used in most monochromators for the visible region. The Wadsworth mounting, shown at *C* and *D*, is sometimes used. It consists of a prism

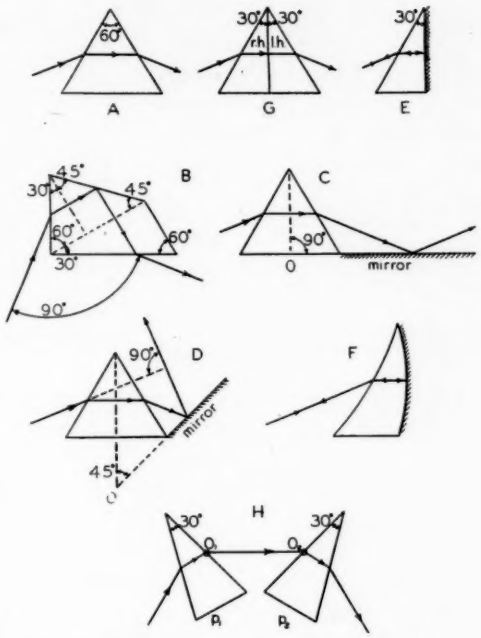


FIG. 21. Various forms and arrangements of dispersing prisms.

and a plane mirror so disposed that they rotate as a unit about an axis *O* at the intersection of the bisector of the refracting angle of the prism with the plane of the mirror. The deviation is constant and can be made of any predetermined value; in *C* it is 0 degree and at *D*, 90 degrees. By combining prism and mirror as shown at *E*, the prism used in the Littrow spectrograph is obtained. The Littrow system is autocollimating, which results in placing the entrance slit and the photographic plate or ocular close together. Moreover, it is difficult to stop out stray light. These inconveniences are less pronounced in very large spectrographs, and the economy of prism material has led to wide use of the Littrow system in such instruments. The Féry system sketched at *F* is similar except that no lenses are required, the prism surfaces being ground spherical instead of plane. When dispersion prisms are made of quartz, certain modifications of the basic forms are often required. Thus the 60° prism shown at *A* must be built up of two 30° prisms of right- and left-handed quartz; the result is the Cornu prism shown at *G*. This prism is used in the

[559] C. M. Tuttle, J. Opt. Soc. Am. **26**, 282–292 (1936).

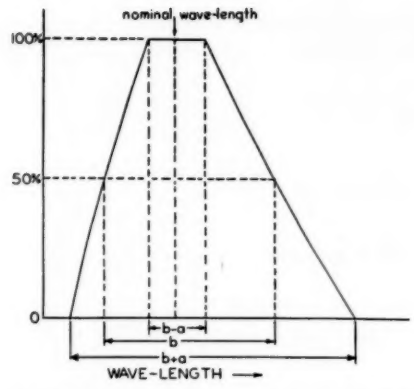


FIG. 22. Spectral transmittance of monochromator with finite slit widths, a and b measured in units of wave-length intervals.

Wadsworth system for the ultraviolet. Owing to the reversal of path through the Littrow and Féry prisms, they need not be composite. A type that is especially suited to quartz because it consists of two symmetrical elements and is economical of material in comparison with its aperture is the Thollon type sketched at H. The two 30° prisms P_1 and P_2 rotate by equal amounts in opposite directions on axes O_1 and O_2, respectively. Diffraction gratings may be used in place of prisms in all of these systems, and also in other arrangements particularly designed for diffraction gratings.

Dispersing systems should be arranged so that aperture remains constant as the wave-length adjustment is varied. For one thing, this means that the prism must not cut into the beam at any position. For another, all the spectral components of the beam must fill the aperture. In other words, the source must effectively subtend a larger angle at the entrance slit than the angular aperture of the optical system, for all wave-lengths. When the source itself does not do this, the remedy is to focus the source on or near the slit with a condenser of adequate aperture.[560]

When the source gives a continuous spectrum, as is usually the case in colorimetry, the energy transmitted by a monochromator cannot be spectrally pure, for that would involve the use of infinitely narrow slits. The formulas for computing the spectral distribution from a monochromator are too complex to discuss here, but monochromators are used so frequently that a typical "spectral transmittance" curve for a monochromator has been sketched in Fig. 22. The widths of the entrance and exit slits have been denoted by a and b, respectively, in terms of the corresponding wave-length intervals.[561] The region of complete transmission (exclusive of instrumental losses) is included in a wave-length interval of $b-a$, and the transmittance is zero outside the range $a+b$. The roles of a and b should be interchanged when the exit slit is narrower than the entrance slit. The extent of the spectral impurity of the energy isolated by a monochromator is indicated by Fig. 22.

If the entire spectrum from a monochromator is passed through another monochromator arranged symmetrically, the result will be an undispersed beam having the same spectral composition as the beam entering the first monochromator (exclusive of reflection and absorption losses). By inserting a template in the plane of the spectrum formed by the first monochromator, the spectral composition of the emergent beam can be altered at will, which makes this form of a double monochromator the equivalent of a filter of which the spectral transmittance curve can be assigned any desired form. Most commonly, a very narrow portion of the spectrum is allowed to pass, the first monochromator serving merely as a filter to furnish the second monochromator with nearly homogeneous energy from which the impurities are almost completely eliminated in the second monochromator. An alternative plan is to reverse the orientation of the second dispersing system, in which case the dispersion is twice that of a single monochromator, instead of being zero. The resultant variation of wave-length across the width of the final slit is objectionable in some applications.

The slit is one of the most important parts of any dispersing instrument. For accurate results, the jaws of the slit should be clean, parallel, sharp, and coplanar. A piece of silvered glass or quartz with a slit ruled in the silver is excellent if no width adjustment is necessary. Adjustable slits may be either unilateral or bilateral. The latter type is required if the center of the slit

[560] D. C. Stockbarger and L. Burns, J. Opt. Soc. Am. **23**, 379–385 (1933).

[561] T. R. Hogness, Z. P. Zscheile, and A. E. Sidwell, J. Phys. Chem. **41**, 379–415 (1937).

image is to remain fixed as the slit width is varied, but the mechanical arrangements are more complicated than for the unilateral slit. Inasmuch as the rays from the ends of the slit traverse a longer path in the prism than does the axial ray, they are more deviated and the spectral lines are therefore curved. This curvature is often neutralized by suitably curving the jaws of the entrance slit. In double monochromators, the exit slit is also frequently curved to secure the maximum efficiency and purity of the energy isolated.

In the visible region, lenses can be so well achromatized that a single focus setting will serve for the entire spectrum. Quartz is not readily achromatized, however, so when the collimator and telescope lenses are of quartz, it is customary to provide a focusing adjustment calibrated according to wave-length. For the same reason, the plate holders of quartz spectrographs are tilted sharply so that the end receiving the shortest wave-lengths is nearer the camera lens. The field is also noticeably curved, and therefore especially thin plates are used and forced into a suitable curvature in the plate holder.

28. VISUAL SPECTROPHOTOMETERS

Essentially, any spectrophotometer consists merely of a photometer associated with a dispersing system, so that relative emittance, transmittance, or reflectance may be measured for any single short spectral region. Many instruments have been designed for this purpose, but only a few of those in common use and ordinarily obtainable can be mentioned here. For more complete surveys of spectrophotometers and more detailed descriptions of the instruments mentioned here, reference should be made to the 1925 Report of the Committee on Spectrophotometry of the Optical Society of America,[1] and to later surveys by the chairman of that committee.[523, 562]

The Lummer-Brodhun spectrophotometer is a clear illustration of the essentials of any spectrophotometer. In this instrument, a Lummer-Brodhun photometer cube is placed near a dispersing prism, as shown in Fig. 23.* Two collimators at right angles project beams into this cube, the beams being respectively transmitted and reflected by the cube into the dispersing prism. A monochromatic photometric field of any type obtainable with a Lummer-Brodhun photometer can be obtained by observing the face of the dispersing prism through the exit slit. The adjustment of the width of the slit in one of the collimators has often been used to vary the luminance of one-half of the field to secure a photometric balance. The limitations of this method of photometric adjustment were suggested in Section 24. This method is even less desirable in spectrophotometry because of the variations of effective wave-length, radiant purity, and resolving power which result from variations of slit width. Brodhun arranged a pair of rotating prisms which caused one of the incident beams to sweep rapidly over the surface of a stationary sector disk, the aperture of which could be varied continuously. Hyde[540] designed a disk with curved sectors so that the average transmittance of the sectors, when placed in front of the slit of one of the collimators in the Lummer-Brodhun spectrophotometer, depends upon the distance between the slit and the center of rotation of the disk. The photometric adjustment is made by moving the center of rotation of the disk, and the relative transmittances of the

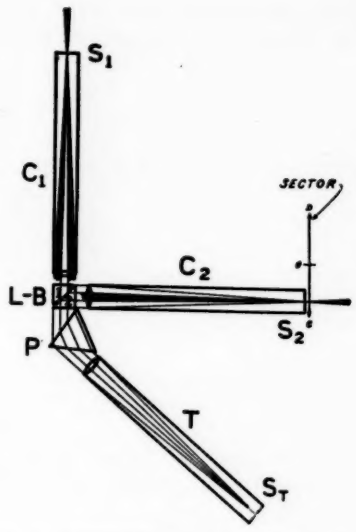


FIG. 23. Lummer-Brodhun spectrophotometer.

[562] K. S. Gibson, J. Opt. Soc. Am. **24**, 234–249 (1934).

* Reproduced from *Measurement of Radiant Energy*, W. E. Forsythe, Editor, by permission of McGraw-Hill Book Company, New York.

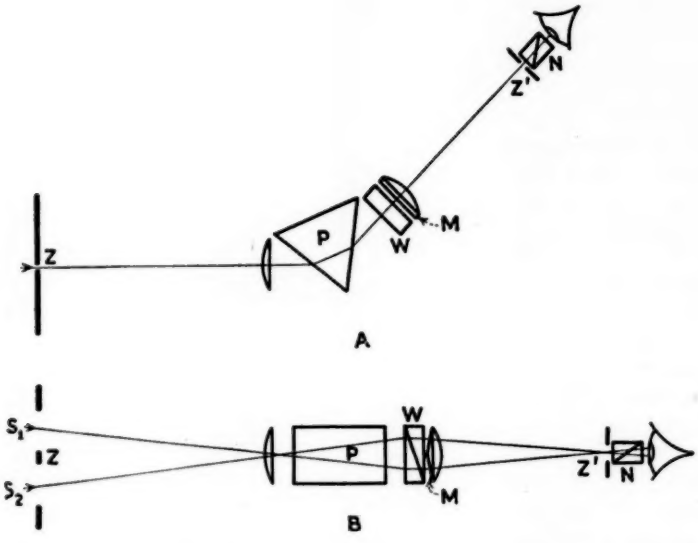


FIG. 24. The optical system of the König-Martens spectrophotometer. A side elevation is shown above, and a development in plan view is shown below.

two collimators are calibrated in terms of the position of this axis. This arrangement is employed in the model made by Schmidt and Haensch. The Lummer-Brodhun spectrophometer is convenient for comparison of the spectral emittances of sources because of the widely separated incident beams, and it has been used primarily for that purpose.

The Keuffel and Esser color analyzer is similar in principle to the Lummer-Brodhun spectrophotometer. The Keuffel and Esser instrument employs a single collimator with a divided slit, before which is placed a special sector disk designed so that the relative transmittance of the sector for the beams entering the two halves of the slit can be varied from zero to 110 percent while the disk is rotating. The sample and the standard are placed in the beams incident on the respective halves of the entrance slit. Images of these two halves are formed on the exit slit by the telescope lens, and these are made coincident by a biprism placed near the dispering prism of a constant-deviation spectrometer. The faces of the biprism observed through the exit slit form juxtaposed photometric fields for one spectral region at a time.

Most other visual spectrophotometers in current use employ polarization photometric devices. These instruments can be divided into two classes, the first having monochromatic photometric fields of the same character as in the Lummer-Brodhun and the Keuffel and Esser spectrophotometers, and the second having photometric fields consisting of portions of juxtaposed spectra, which therefore exhibit variations of hue along the photometric boundary. Non-polarizing spectrophotometers employing juxtaposed spectra are conceivable, but none is in common use at present. A somewhat similar principle is employed in photographic spectrophotometry.[563]

The Gaertner spectrophotometer consists of a constant-deviation spectrometer, a Martens polarizing photometer, and accessories for holding and illuminating samples and standards. An image of the photometric field of the Martens photometer is focused through the entrance slit on the collimator lens of the spectrometer. The telescope lens is observed through the exit slit, without an eye lens, and a monochromatic photometric field appears in the dispersing prism, the effective wave-length being determined by the portion of the spectrum which is refracted through the exit slit. This wave-length can be changed, as in the Keuffel and Esser color analyzer, by simply rotating the constant deviation prism.

The König-Martens spectrophotometer (made by Schmidt and Haensch) is similar in principle

[563] H. E. Howe, Phys. Rev. **8**, 674–688 (1916).

to the Gaertner spectrophotometer but is designed and built as a single unit. The simplification of the optical system obtained in this way eliminates much stray light and many aberrations, and a much more sharply defined photometric field results. In the König-Martens spectrophotometer, Fig. 24,* the entrance slit Z is divided in half (as in the Keuffel and Esser color analyzer) and images Z' of these two halves of the entrance slit are focused on the exit slit by the collimator and telescope lenses. These images are made coincident by a combination of a biprism M and a Wollaston prism W placed in contact with the telescope lens. The Wollaston prism supplements the action of the biprism and polarizes (in perpendicular planes) the coincident images of the two halves of the entrance slit. Three other images of each half of the entrance slit which are formed by the combination of the biprism and the Wollaston prism are refracted away from the exit slit and are absorbed or otherwise removed from the optical system. The photometric field formed by the faces of the biprism M can be observed through the exit slit without an eye lens.

The relative luminances of the two halves of the photometric field can be adjusted to balance by rotating a Nicol prism N placed between the exit slit and the eye. The angle between the principal plane of the Nicol and the plane of polarization of the beam from one of the halves of the photometric field can be read from a divided circle. The spectral transmittance or reflectance of a sample placed in the beam incident on one of the halves of the entrance slit can be calculated by use of the formula for the Martens photometer, discussed in Sections 13 and 26. In most polarizing spectrophotometers a balance of 45 degrees cannot be obtained with unobstructed beams, because of polarization introduced by oblique reflections from the dispersing prism and other optical parts. This effect must be taken into consideration in computing the transmittance (or reflectance relative to a standard) by use of the tangent-square rule. If θ_0 is the angle required for match when no absorbing samples are in either beam, and if θ_1 is the angle required for match when an absorbing sample is in the sample beam, the transmittance (or relative reflectance) of the sample is

$$t_\lambda(\text{or } r_\lambda) = \cot^2 \theta_0 \cdot \tan^2 \theta_1.$$

Errors caused by partial polarization by the sample can be largely eliminated by placing the sample first in the sample beam and then in the standard beam. If the angles required (at wavelength λ) for match in these two cases are θ_1 and θ_2, the spectral transmittance (or relative reflectance) is

$$t_\lambda(\text{or } r_\lambda) = \tan \theta_1 \cdot \cot \theta_2.$$

Like the Gaertner spectrophotometer, the Bausch and Lomb spectrophotometer consists of an assemblage of a constant deviation spectrometer, a Martens photometer, and auxiliary apparatus for holding and illuminating the sample. In the Bausch and Lomb spectrophotometer, however, an image of the photometric field of the Martens photometer is formed on the entrance slit, with the dividing line perpendicular to the length of the slit. The exit slit is examined with an ocular lens, and this slit is opened as wide as is necessary to provide a satisfactory photometric field. This field consists of two juxtaposed spectra. Noticeable differences of hue can be detected along the line separating the two spectra. Considerable variations of luminance can also occur along the dividing line, especially when materials with rapidly changing transmittances (or reflectances) are measured. Such non-uniformities in the photometric field force a reduction of the width of the exit slit and reduce the precision of photometric adjustment. Such non-uniformities, however, assist in the detection of very narrow absorption bands, which frequently remain undetected in spectrophotometers employing monochromatic (but not spectrally homogeneous) photometric fields. Stray light, which is responsible for the most serious errors in the use of spectrophotometers with monochromatic fields, is distributed equally over both fields in most instruments employing juxtaposed spectra, and the only errors arising from this stray light are the result of decreased sensibility. In general, the photometric dividing line cannot be made to disappear as completely in spectrophotometers employing juxtaposed spectra as in those employing monochromatic fields.

* Reproduced from *Principles of Optics*, A. C. Hardy and F. H. Perrin, by permission of McGraw-Hill Book Company, New York.

The Hüfner spectrophotometer, built by Hilger, consists of a rhomb placed directly in front of the entrance slit of a constant-deviation spectrometer. This rhomb illuminates, with a sharp dividing line, the two halves of the entrance slit with light from two separated but parallel beams from the same source. A transmitting sample may be placed in one of the beams. The other beam is polarized by passage through a fixed Nicol prism. The dividing edge of the rhomb is so close to the entrance slit that it is in good focus in the plane of the exit slit, and the two juxtaposed spectra that are formed there are observed through an eye lens. An analyzing Nicol prism in a divided circle is placed between the telescope lens and the exit slit. Rotation of this prism affects the luminance of only the portion of the photometric field formed by the polarized incident beam, and does not influence the spectrum formed by the beam transmitted by the sample. When properly designed and adjusted, the rhomb of the Hüfner spectrophotometer completely compensates for the polarization by reflection at the surfaces of the dispersing prism, and the transmittance of the sample is the square of the sine of the angle of rotation of the analyzing Nicol from the extinction position.

The Hilger-Nutting spectrophotometer is similar in principle and operation to the Bausch and Lomb spectrophotometer, the Martens photometer of the latter being replaced with a photometer unit consisting of two fixed polarizing Nicol prisms, a rhomb for uniting two parallel beams (more widely separated than in the Bausch and Lomb or Hüfner spectrophotometers), and an analyzing Nicol in a divided circle. An image of the uniting edge of the rhomb is focused on the center of the slit of the spectrophotometer, and juxtaposed spectra are observed by an eye lens focused on the exit slit. The transmittance of the sample is proportional to the square of the tangent of rotation of the analyzing Nicol from extinction of the comparison beam.

Spectrophotometric measurements can be checked with an ordinary photometer, by use of a source radiating a single spectral line. Such sources are rare, but filters are available which isolate single lines from several sources and provide means of irradiation with essentially homogeneous energy.[564]

Light sources with continuous spectra are employed in abridged spectrophotometry. In this method the reflectance or transmittance of a sample is measured when each of a series of "narrow band" filters is placed in the beam, either between the source and the sample or between the sample and the photometer. An approximate average transmittance or reflectance is obtained in this way for each of the rather broad bands of wave-lengths transmitted by the auxiliary filters. The average effective wave-length for each selection filter is difficult to define in any useful way. As mentioned in Section 11, transmittances measured with broad bands such as are employed in abridged spectrophotometry do not obey either Bouguer's or Beer's laws except in the case of nearly non-selective samples. Because of the simplicity of the apparatus and method, abridged spectrophotometry is frequently useful in routine measurements, especially where highly selective samples are not encountered.[565]

29. PHOTOELECTRIC SPECTROPHOTOMETERS

The difficulties of making accurate visual measurements of spectral transmittance, and especially of spectral reflectance, near the ends of the visible spectrum have stimulated the development of numerous photoelectric spectrophotometers.

In the Cenco spectrophotelometer a concave diffraction grating is employed in the monochromator. Provision is made for interchanging equal thicknesses of solution and solvent in the beam emerging from the monochromator. The transmitted beam is incident upon a barrier layer photoelectric cell and the response of this cell is indicated by a sensitive galvanometer. The spectral transmittancy of a solution is computed as the ratio of the galvanometer readings when the solution and solvent are placed before the photo-cell.

The Coleman model 11 spectrophotometer is essentially the same as the Cenco instrument except that a plane diffraction grating is em-

[564] K. S. Gibson, J. Opt. Soc. Am. **25**, 131–135 (1935).
[565] J. A. Van den Akker, Paper Trade J. (technical section) **111**, 142–151 (1940).

ployed in the monochromator and the source rather than the grating is moved in order to select various wave-lengths.

A double monochromator employing two diffraction gratings is incorporated in the Coleman model 10 spectrophotometer which was designed by R. W. Wood. This instrument employs a caesium-oxide photoemission tube, and the photo-cell current is measured by use of a potentiometer.

The Beckman spectrophotometer employs a quartz prism in a Littrow mounting with a concave focusing mirror, so that measurements can be made throughout the wave-length range from 200 mμ to 2000 mμ. Several electron emission photo-cells are used with amplification to cover this range. Gibb[566] describes these and other photoelectric spectrophotometers, with diagrams, and a discussion of the use of some of these instruments, with photographs, was published by Müller.[567] A photoelectric spectrophotometer incorporating many unique features has been built and described by Harrison.[568]

The spectrophotometry of opaque reflecting substances and especially of near-white materials, for which visual spectrophotometers are inadequate, can be accomplished satisfactorily by only the most carefully designed photoelectric spectrophotometers. The General Electric recording photoelectric spectrophotometer, designed by A. C. Hardy specifically for the measurement of spectral reflectance, can also be used without modification for the measurement of spectral transmittance. This instrument employs a double monochromator for the elimination of stray energy. The energy supplied by this monochromator passes through a polarizing photometer somewhat similar in principle to the Martens photometer. An additional polarizing prism is rotated at high speed to produce alternately flickering beams which are reflected from the sample and the standard and ultimately fall upon the photoelectric cell. The fluctuations in the photo-cell current resulting from inequalities in the reflected intensities of these beams are amplified and actuate a motor that rotates the balancing prism of the photometer. Through a mechanical coupling, the reflectance corresponding to the position of balance is recorded on a graph sheet, which is rotated continuously under the recording pen as the wave-length changes.[569–573]

30. SPECTROPHOTOMETRY

The results of spectrophotometric measurements depend upon the control of several variables that do not appear in the ordinary report of data. The importance of spatial conditions of irradiation and collection which are discussed in Sections 5 and 10 applies to an equal extent to spectrophotometry. The conditions adopted should be indicated in all reports of data. The International Commission on Illumination, in 1931, recommended irradiation at 45 degrees to the surface being tested and unidirectional collection normal to the plane of the surface.[505] Although conditions similar to these (supplemented with tolerance limits necessary for the practical realization of these conditions) are desirable in principle, and indeed have been employed in some instruments, such conditions have not proved to be practical for photoelectric spectrophotometry because of the relative inefficiency of optical systems designed in accordance with these specifications. In the General Electric instrument, the incident beam is normal to the surface of the sample, and over 99 percent of the diffusely reflected energy is collected with uniform efficiency. Unfortunately, some indefinite fraction of the specularly reflected energy is also collected. This is serious, however, only when glossy samples are measured, and the indefiniteness of these conditions of collection can be remedied by a slight modification of the collecting sphere. This modification is such that the incident energy strikes the sample within a cone of less than two degrees diameter, with the central ray six degrees from the normal to the surface of the sample, and the specularly reflected

[566] Thomas R. P. Gibb, Jr., *Optical Methods of Chemical Analysis* (McGraw-Hill, New York, 1942), Chapter II.
[567] R. H. Müller, Ind. Eng. Chem. Anal. Ed. **13**, 667–754 (1941).
[568] G. R. Harrison and E. P. Bentley, J. Opt. Soc. Am. **30**, 290–294 (1940).
[569] A. C. Hardy, J. Opt. Soc. Am. **25**, 305–311 (1935).
[570] A. C. Hardy, J. Opt. Soc. Am. **28**, 360–364 (1938).
[571] J. L. Michaelson, J. Opt. Soc. Am. **28**, 365–371 (1938).
[572] K. S. Gibson and H. J. Keegan, J. Opt. Soc. Am. **28**, 372–385 (1938).
[573] O. W. Pineo, J. Opt. Soc. Am. **30**, 276–289 (1940).

energy can be completely collected or completely eliminated at will.

Definiteness in specifying the mode of irradiation and of collection is more desirable in practical spectrophotometry than is an indefinite approximation to theoretically ideal conditions. Measurements of glossy samples are most subject to errors arising from variations of conditions of irradiation and collection and nearly perfect diffusing samples are least subject to such errors. Magnesium oxide deposited on a smooth surface in the form of smoke from a burning ribbon or pile of chips is not only an almost perfect diffuser, but is also the most perfect reflector that can be made in a reproducible manner. Consequently, such a surface of magnesium oxide is widely used and recommended as a standard for the spectrophotometry of reflecting materials. Some laboratories employ working standards of more durable materials, which can be cleaned, but such surfaces are calibrated by comparison with magnesium oxide to guard against slow changes in the reflectance of the permanent standard. Working standards with glossy surfaces should be avoided if the spatial conditions of irradiation or collection are subject to the slightest variations.

When transmission measurements are being made, the effects of spatial distribution of the incident and collected energy are less frequently important than in reflectance measurements. Rays oblique to the surface, however, have paths greater than the thickness of the sample, and are consequently more strongly absorbed than normally incident beams. The focusing effects of samples, even those with plane parallel surfaces, sometimes disturb the optical characteristics of a spectrophotometer and compensation or correction is required. The effect of scattering in transmittance measurements, discussed in Section 5, is equally important in spectrophotometry, and spectral transmittance data for scattering samples should be supplemented with specifications of the spatial conditions of irradiation and collection.

In all cases, for spectrophotometric measurement of both transmitting and reflecting samples, the conditions of irradiation and collection should be identical for the sample and the standard with which it is compared. In systems employing unidirectional irradiation, the sample and the standard may be placed successively in one beam and compared with an undisturbed comparison beam. For any mode of irradiation beams may be received simultaneously from the sample and standard and directly compared by use of equivalent collection systems. Such a method is required when the samples are diffusely irradiated, which type of irradiation can be used if the sample and the standard simultaneously occupy openings in a single diffusing sphere. In photoelectric spectrophotometers employing alternately flickering beams (as in the General Electric instrument), the same collecting system can be used for both the sample and the standard, fulfilling the requirement of equivalence. These methods, of course, are all special applications of the substitution method of photometry.

Errors in spectrophotometry often arise from the effects of polarization mentioned in Section 16. Such errors can be detected and eliminated by a consideration of the facts discussed in Section 16.

The techniques of wave-length calibration, the detection and elimination of stray light, and the correction of errors due to the finite slit widths employed in spectrophotometers are described in some detail in articles devoted primarily to spectrophotometry.[540, 562] For that reason, they will not be discussed here, but it is to be emphasized that they are of the utmost importance in accurate spectrophotometry. Color specification based on spectrophotometric data demands the most accurate technique of measurement. *The chromatic sensibility of normal human vision corresponds to a degree of accuracy that is rarely surpassed in spectrophotometry and that can be attained only by the most painstaking technique employing the most carefully designed and constructed apparatus.*

A Graphic Calculator for Spectrochemical Analyses

A. P. VANSELOW AND GEORGE F. LIEBIG, JR.
University of California Citrus Experiment Station, Riverside, California
(Received January 10, 1944)

ANYONE who has had an opportunity to use a graphic calculator for computing quantitative spectrographic analyses realizes the enormous convenience of such a device, the chief advantages being a great saving of time and a greatly reduced probability of arithmetical error. The calculator herein described is the result of combining what were considered the better features of several calculators described in the literature[1] with our own ideas. The design and construction of the device were influenced by such factors as the following: the type of spectrographic work being done in this laboratory, the materials available, and rather limited shop facilities.

A view of the calculator is shown in Fig. 1. In brief, it consists of a box housing a roll of white paper, which passes out over a drawing board on top of the box and back into the box onto another roller. The top of the box has on it a transparent slider, with its vertical hairline and a logarithmic scale of percentage transmission, and a sliding holder for the charts of analytical scales of the percentages of the various elements. There is also on the box top a scale of the relative intensities of the iron lines that are used for the film or plate calibration.

The advantages of this calculator are: No expensive cross section or logarithm paper is required other than that used in preparing the customary working curves; the percentage transmission scale engraved adjacent to the vertical hairline on the transparent slider is always easily read and eliminates the necessity of scanning all the way across to the edge of the paper for the scale numbers. Having the transmission or blackening scale engraved on the transparent slider also eliminates the need of writing the scale numbers on the calibration curve each time a new curve is drawn—a practice followed by some workers.

The calculator box is constructed of wood and is 18 inches (45.7 cm) wide, 22 inches (55.9 cm) long, 9 inches (22.9 cm) deep in front, and 12.5 inches (31.8 cm) deep in back. This box contains a roll of white drug bond paper, 15 inches (38.1 cm) wide. The paper passes out over a removable drawing board, 11.25 inches (28.6 cm) wide, and back into the box onto a roller. Each roller is provided with a crank on the outside of the box, to aid in keeping the paper taut and the film calibration curve properly aligned. The rollers are 1-inch (2.54-cm) maple dowels, the one receiving the used paper being slotted similarly to the receiving spool in a roll-film camera. The axes of the two rollers and the edges of the drawing board must be parallel to avoid wrinkles in the paper on the drawing board. Instead of being wound onto the second roller, the used paper could, of course, be pulled out through a slot, torn off, and discarded; the authors prefer to use the two rollers, since this arrangement provides a simple and effective way of keeping the paper taut and easily adjusted as to the proper position of the calibration curves.

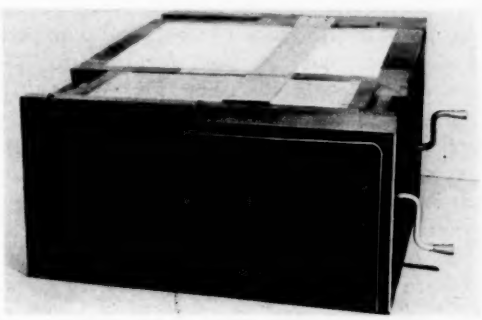


FIG. 1. Graphical calculator—front view.

[1] C. King, J. Opt. Soc. Am. **32**, 112 (1942); R. H. Müller, Ind. Eng. Chem. Anal. Ed. **13**, 667 (1941); J. S. Owens, *Proceedings of Fifth Summer Conference on Spectroscopy and Its Application* (John Wiley & Sons, Inc., New York, 1938), p. 17; J. S. Owens, Ind. Eng. Chem. Anal. Ed. **10**, 64 (1938); A. M. Sampson, *Proceedings of Fifth Summer Conference on Spectroscopy and Its Application* (John Wiley & Sons, Inc., New York, 1938), p. 8.

The inset scale (Fig. 2, *A*) is that of the relative intensities of the iron lines used for the calibration of the film. These relative intensities are plotted on a logarithmic scale of a 3-inch (7.6-cm) cycle. A 3-inch cycle may seem rather small, but the authors find that errors involved in other phases of their spectrochemical methods do not warrant the use of a more extended scale. The wave-lengths and the corresponding relative intensities of the iron lines used for the film calibration are: 3163.9, 0.19; 3165.0, 0.42; 3165.9, 0.67; 3166.4, 1.00; 3168.9, 0.31; 3175.4, 2.00; 3180.2, 3.63; and 3196.9, 6.3. These relative intensities differ slightly from the values furnished by Hasler[2] and by Hasler and Harvey,[3] but our iron spectra are produced in a manner somewhat different from theirs.

The standard iron spectra are made by using an iron No. 4 Gimp tack held in a 0.0625-inch (1.59-mm) hole in a 0.25-inch (0.635-cm) graphite rod. The upper electrode of graphite is 0.125 inch (3.18 mm) in diameter; the electrode separation is 0.158 inch (4 mm); and 6 amperes of direct current are used for the arc. The exposure is made by using four consecutive arcings of 5 seconds' duration, separated by cooling-off periods of 5 seconds. The lines of the iron spectra have shown surprisingly constant relative intensities. As many as five such iron spectra (on one film), having relative exposures of 1, 2, 4, 8, and 16, have been found, when measured on the ARL densitometer, to satisfy the "self-consistency" test of Harrison.[4] Even though all of the eight iron lines used by the authors may not satisfy the requirements as set forth by Dieke and Crosswhite,[5] the production of the iron spectra is evidently sufficiently uniform to give satisfactory constancy of the relative intensities for these iron lines.

Only one set of iron lines is used for the calibration of the film. We have found only minor variations of gamma for the different wave-length regions with the Eastman Spectrum No. 2 film when developed to a gamma of 1.2. The data of Hasler and Harvey[3] for the same film also fail to show any great variation of gamma with wave-length, although in Hasler's work,[2] presumably with the same film, there is considerable variation of gamma with wave-length. Be that as it may, we elected to sacrifice some measure of accuracy for the sake of simplicity. We are using some iron lines of considerably greater intensity than the strongest used by Hasler and Harvey. We do this because we have found that when the exposure is restricted to keep the background at a minimum, some of the weaker lines in their list may not register on the film. We have, however, found no evidence that these stronger iron lines are subject to reversal effects. We have found the iron spectrum, as suggested by Slavin[6] and recommended by Hasler,[2] a simple and effective method of film calibration.

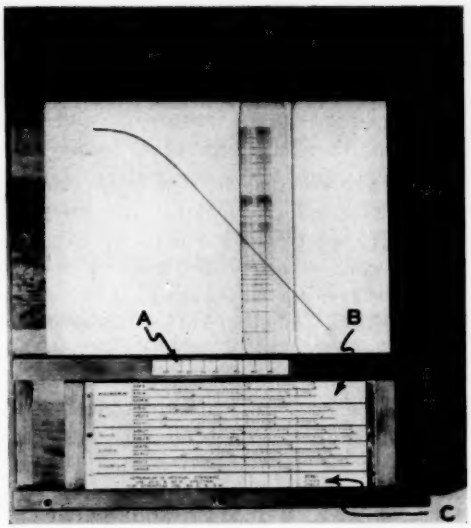


FIG. 2. Graphical calculator—top view: *A*, relative intensity scale of iron lines; *B*, percentage scales of the elements; *C*, indices for internal-standard lines.

The vertical hairline slider, of Lucite, is 20 inches (50.8 cm) long, 2.25 inches (5.7 cm) wide, and 0.125 inch (0.32 cm) thick. It is securely fastened at right angles to the cross-head, which is made of maple and slides freely but not

[2] M. F. Hasler, J. Opt. Soc. Am. **31**, 140 (1941).
[3] M. F. Hasler and C. E. Harvey, Ind. End. Chem. Anal. Ed. **13**, 540 (1941).
[4] G. R. Harrison, J. Opt. Soc. Am. **18**, 492 (1929); H. H. Grossman, R. A. Sawyer, and H. B. Vincent, J. Opt. Soc. Am. **33**, 185 (1943); R. A. Sawyer and H. B. Vincent, J. Opt. Soc. Am. **33**, 247 (1943).
[5] G. H. Dieke and H. M. Crosswhite, J. Opt. Soc. Am. **33**, 425 (1943).
[6] Morris Slavin, Ind. Eng. Chem. Anal. Ed. **12**, 131 (1940).

loosely in the close-fitting maple guides. A small amount of paraffin on the sliding surfaces acts as a lubricant. The back surface of the Lucite slider has engraved on it the percentage transmission plotted on a logarithmic scale with a 3-inch (7.6-cm) cycle. The vertical index line, which extends the full length of the slider, and the logarithmic graduations were engraved with a phonograph recording needle. The numbers were engraved by means of a specially prepared set of stencils and a stylus. The engravings were made visible by filling them in with India ink, the excess ink being easily removed, after drying, by wiping with a slightly dampened cloth. The logarithm of the percentage transmission is the scale used on the slider because the calculator is used in conjunction with the data obtained on the ARL densitometer, which instrument gives the transmission of the spectrum line in percent of the transmission of the clear film. Of course, any convenient scale of density, galvanometer deflection, or blackening can be used in place of the scale used by the authors.

The method of preparing the analytical scales for the percentages of the elements is described by Owens.[1] The customary working curves of the percentage of element against the intensity ratios were plotted on 3-inch (7.6-cm) cycle log-log paper; from these curves the analytical scales were prepared. Close inspection of Fig. 2 shows that several spectrum lines are used for each element and three lines of the internal standard. When the analytical scales for several elements with which the same internal standard is used are plotted on one card, care must be taken to have the internal-standard indices of all scales coincide. If this is done, one setting of the scales will suffice for all the elements on the card. The cards of the analytical scales fit into place under two thin, narrow strips of Lucite, attached near the ends of the sliding board (Fig. 2, *B*). The card holder is 5 inches (12.7 cm) wide and 14.5 inches (36.8 cm) long. A strip of felt is glued to each of the two sliding edges of the holder, so that it has a snug sliding fit between the maple guides; enough friction is wanted to insure against accidental displacement of the scales.

The manner of using a graphic calculator is adequately described by King[1] and Owens;[1] there is no need of going into detail here regarding its use. There is, however, one point about which much could be said and which should be emphasized more than it usually is. There is no simple way of correcting for background when using a graphic calculator; it is highly important, therefore, to control the exposure properly, so as to avoid the production of any appreciable background. Much has been written concerning the correction for background;[2,3,7] the usual method of correction is to subtract the intensity of the background from that of the line, both intensities being derived from the film calibration curve. This method seems reasonably sound, but the findings of Webb[8] certainly cast doubt on its universal validity. It was found in this laboratory that the total effect of background and line depended upon whether the background was put on before, simultaneously with, or after the line image. These results were quite disturbing, and it was highly gratifying later to come across this work of Webb concerning the effect of successive exposures of different intensity levels. Thus it seems that the mode of correction depends upon whether or not line image and background are produced simultaneously, and perhaps also upon their relative intensities. Because of this, it is again emphasized that the best procedure is to regulate the exposure so that background is eliminated as completely as is possible.

---

[7] J. Cholak and R. V. Story, J. Opt. Soc. Am. **32**, 502 (1942); W. C. Pierce and N. H. Nachtrieb, Ind. Eng. Chem. Anal. Ed. **13**, 774 (1941); L. W. Strock, *Proceedings of Seventh Summer Conference on Spectroscopy and Its Applications* (John Wiley & Sons, Inc., New York, 1940), p. 134; L. W. Strock, J. Opt. Soc. Am. **32**, 103 (1942).

[8] J. H. Webb, *Proceedings of Sixth Summer Conference on Spectroscopy and Its Application* (John Wiley & Sons, Inc., New York, 1939), p. 157. Paper No. 500, University of California Citrus Experiment Station, Riverside, California.

Impact of a Wave Packet and an Absorbing Particle

HERBERT E. IVES
Bell Telephone Laboratories, Inc., New York, New York
(Received January 20, 1944)

INTRODUCTION

THE impact of a wave packet on an absorbing particle presents a more complicated problem than the previously treated case of impact on a reflecting particle.[1] After impact the radiation no longer exists as such, and it becomes necessary to study the characteristics of the form of energy into which the radiant energy is transformed. This may be done by noting that the absorption of radiant energy results in the production of heat, which is kinetic energy of the atoms or subparticles of which the absorbing particle is composed. Then by considering the change of mass of these atoms due to their motion, the over-all change of mass of the absorbing particle can be obtained, and the equations of motion derived.

(1) STATIONARY ABSORBING PARTICLE

We start by setting up the expression for the momentum of a particle moving as a whole with velocity v, in which, because of absorbed heat, the constituent atoms are in random motion. The total momentum $\bar{M}$ will be

$$\bar{M}=\frac{\sum m(v\pm w)}{\left[1-\frac{(v\pm w)^2}{c^2}\right]^{\frac{1}{2}}}, \tag{1}$$

where m is the rest mass of one of the constituent atoms and $v\pm w$ is its velocity of heat motion; the denominator on the right is in accord with the variation of mass with motion derived in the preceding paper.[1] The symbol w in this equation is as yet of undefined value, expressing merely the fact that the atoms have velocities additional to the velocity v of the particle as a whole.

We now for simplicity consider the atoms in two groups, those moving, with respect to the center of the particle, in the direction of motion of the particle and those against, and treat these as all moving with uniform velocities in each group. We designate these two velocities as $v+w_1$ and $v-w_2$. We then rewrite (1) as

$$\bar{M}=\frac{1}{2}\sum\frac{m(v+w_1)}{\left[1-\frac{(v+w_1)^2}{c^2}\right]^{\frac{1}{2}}}+\frac{1}{2}\sum\frac{m(v-w_2)}{\left[1-\frac{(v-w_2)^2}{c^2}\right]^{\frac{1}{2}}} \tag{2}$$

which we can break down into

$$\bar{M}=\frac{1}{2}\sum\frac{mv}{\left[1-\frac{(v+w_1)^2}{c^2}\right]^{\frac{1}{2}}}+\frac{1}{2}\sum\frac{mv}{\left[1-\frac{(v-w_2)^2}{c^2}\right]^{\frac{1}{2}}}+\frac{1}{2}\sum\frac{mw_1}{\left[1-\frac{(v+w_1)^2}{c^2}\right]^{\frac{1}{2}}}-\frac{1}{2}\sum\frac{mw_2}{\left[1-\frac{(v-w_2)^2}{c^2}\right]^{\frac{1}{2}}}. \tag{3}$$

Now if the particle as a whole is to move with the velocity v, the summation of the momenta of the atoms forward and back, with reference to the center of motion of the particle, must be zero, that is, the momentum described by the last two terms of (3) must be zero, or

$$\frac{w_1}{\left[1-\frac{(v+w_1)^2}{c^2}\right]^{\frac{1}{2}}}=\frac{w_2}{\left[1-\frac{(v-w_2)^2}{c^2}\right]^{\frac{1}{2}}}. \tag{4}$$

[1] H. E. Ives, J. Opt. Soc. Am. **33**, 163–166 (1943).

We now seek the values of w_1 and w_2. We note that if the particle is brought to rest, without escape of any of the heat, the heat motions in the stationary particle will be equal in all directions, their velocities assuming the value w_0, so that $w_1=w_2=w_0$. We can accordingly put

$$w_1=w_0f_1(w_0, v),$$
$$w_2=w_0f_2(w_0, v), \tag{5}$$

f_1 and f_2 being functions which, when $v=0$, have the value unity. Rewriting (4)

$$\frac{w_0f_1(w_0, v)}{\left[1-\left(\frac{v+w_0f_1(w_0, v)}{c}\right)^2\right]^{\frac{1}{2}}} = \frac{w_0f_2(w_0, v)}{\left[1-\left(\frac{v-w_0f_2(w_0, v)}{c}\right)^2\right]^{\frac{1}{2}}}. \tag{6}$$

Replacing $f_1(w_0, v)$ by F_1 and $f_2(w_0, v)$ by F_2, this equation becomes

$$\frac{F_1}{\left[1-\left(\frac{v+w_0F_1}{c}\right)^2\right]^{\frac{1}{2}}}=\frac{F_2}{\left[1-\left(\frac{v-w_0F_2}{c}\right)^2\right]^{\frac{1}{2}}}. \tag{7}$$

By squaring and reduction we get from this

$$\frac{1}{F_1}-\frac{1}{F_2}=\frac{2vw_0}{c^2\left(1-\frac{v^2}{c^2}\right)}. \tag{8}$$

This is satisfied by

$$\frac{1}{F_1}=\frac{Q+\frac{vw_0}{c^2}}{1-\frac{v^2}{c^2}},$$
$$\frac{1}{F_2}=\frac{Q-\frac{vw_0}{c^2}}{1-\frac{v^2}{c^2}}, \tag{9}$$

where Q is undetermined. Substituting these values in (5) we get

$$w_1=\frac{w_0\left(1-\frac{v^2}{c^2}\right)}{Q+\frac{vw_0}{c^2}}=\frac{w_0\left(1-\frac{v^2}{c^2}\right)}{1+\frac{vw_0}{c^2}},$$
$$w_2=\frac{w_0\left(1-\frac{v^2}{c^2}\right)}{Q-\frac{vw_0}{c^2}}=\frac{w_0\left(1-\frac{v^2}{c^2}\right)}{1-\frac{vw_0}{c^2}}, \tag{10}$$

since, when $v=0$, $w_1=w_2=w_0$, the value of Q must be unity.[2]

Inserting these values in (3) we have

$$\bar{M}=\frac{1}{2}\sum\frac{m\left(1+\frac{vw_0}{c^2}\right)v}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}\left[1-\frac{w_0^2}{c^2}\right]^{\frac{1}{2}}} + \frac{1}{2}\sum\frac{m\left(1-\frac{vw_0}{c^2}\right)v}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}\left[1-\frac{w_0^2}{c^2}\right]^{\frac{1}{2}}}+0 \tag{11}$$

or

$$\bar{M}=\sum\frac{mv}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}\left[1-\frac{w_0^2}{c^2}\right]^{\frac{1}{2}}} = \frac{M_0v}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}\left[1-\frac{w_0^2}{c^2}\right]^{\frac{1}{2}}}, \tag{12}$$

where M_0 is the rest mass of the particle.

---

[2] Note that we have incidentally derived, from non-relativistic considerations (for this particular problem), the Einstein formula for the composition of velocities, for if $w_1=\frac{w_0\left(1-\frac{v^2}{c^2}\right)}{1+\frac{vw_0}{c}}$ then $v+w_0=v+\frac{w_0\left(1-\frac{v^2}{c^2}\right)}{1+\frac{vw_0}{c^2}}=\frac{v+w_0}{1+\frac{vw_0}{c^2}}$. Equation (14) derived below has in fact been derived elsewhere [Becker, *Theorie der Electricität* (Teubner, Leipzig, 1933), Vol. 2, p. 348] by assuming that the atoms in the absorbing body move so that their velocities conform to the Einstein formula.

We can rewrite this in the form

$$\bar{M}=\frac{M_0 v}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}+\frac{M_0 v}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}\left[\frac{1}{\left[1-\frac{w_0{}^2}{c^2}\right]^{\frac{1}{2}}}-1\right]. \quad (13)$$

Observing that the quantity in brackets is the kinetic energy of the atoms due to their heat motion[3] divided by M_0c^2, and calling this heat energy $\bar{H}$, we have finally for the momentum of the absorbing particle,

$$\bar{M}=\frac{\left(M_0+\frac{\bar{H}}{c^2}\right)v}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}. \quad (14)$$

From the momentum we can by integration get the kinetic energy of the absorbing particle after the impact, which is

$$\bar{K}=\left(M_0+\frac{\bar{H}}{c^2}\right)c^2\left[\frac{1}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}-1\right]. \quad (15)$$

We now consider the relation between the energy of the wave packet incident on the particle and the energy imparted to the particle. If E_s is the energy density from the (stationary) source at the particle, and Δt the duration of emission by the source, we have

Energy content of wave packet $=E_s(c\Delta t)$.

Now this becomes transformed into the kinetic energy of the particle [Eq. (15)], plus the heat absorbed by the particle. The heat absorbed by the particle in the impact is equal to H, the average rate of heating during the impact, times t, the duration of the impact, or Ht. We have then from the *conservation of energy*

$$E_s(c\Delta t)=\left[M_0+\frac{Ht}{c^2}\right]c^2\left[\frac{1}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}-1\right]+Ht. \quad (16)$$

[3] Strictly speaking it is the kinetic energy when the particle is brought to rest, but since we have postulated that w_0 is the value of velocity when the particle is brought to rest *without escape of heat*, this value of kinetic energy also applies to the particle in motion.

This, by rearrangement, can be written

$$E_s\frac{(c\Delta t)}{c^2}+M_0=\frac{M_0+\frac{Ht}{c^2}}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}, \quad (17)$$

which makes (16) also a statement of the *conservation of mass* upon the acceptance of $E_s[(c\Delta t)/c^2]$ as the mass equivalent of the packet of radiation.

If in the impact we consider $E_s[(c\Delta t)/c^2]$ as a mass, and proceed to the next step of assigning to it a momentum, by forming the product of its mass by its velocity (which is c) we can set up from (14) and (17) the equation for the *conservation of momentum* in the impact, which is

$$E_s\frac{(c\Delta t)}{c}=\frac{\left(M_0+\frac{Ht}{c^2}\right)v}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}. \quad (18)$$

We have now (accepting these assumptions) all the necessary material for finding (1) the velocity attained by the absorbing particle after the impact, and (2) the heat absorbed by the particle in the impact.

To obtain (1) we substitute the value of $[M_0+(Ht/c^2)]/[1-(v^2/c^2)]^{\frac{1}{2}}$ from (18) in (17) which gives

$$E_s\frac{(c\Delta t)}{c^2}+M_0=E\frac{(c\Delta t)}{cv}$$

or

$$v=\frac{E_s(c\Delta t)}{E_s(c\Delta t)+M_0c^2}c. \quad (19)$$

To obtain (2) we substitute the value of $M_0+(Ht/c^2)$ from (18) in (16) which gives

$$E_s(c\Delta t)=E_s\frac{(c\Delta t)}{cv}\times\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}c^2\left[\frac{1}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}-1\right]+Ht \quad (20)$$

or

$$Ht = E_s(c\Delta t)\left[1-\frac{c}{v}\left(1-\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}\right)\right]$$

$$\left(\leqq E_s(c\Delta t)\left[1-\frac{v}{2c}\right]\right).$$

(2) SYSTEM IN UNIFORM MOTION

An important problem now to be investigated is the result to be expected if the impact experiment is done on a platform moving with the uniform velocity V. Specifically, will the end velocity of the particle appear to be the same under this condition, and will the heat absorbed by the particle be the same? The changes introduced by this motion are two in number. First, the energy in the wave packet emitted in the apparent time Δt will be altered. Second, the measurement of velocities will be altered by the changes in clock rates and in measuring rods.

For the first of these changes we note that the emission time becomes $\Delta t/[1-(V^2/c^2)]^{\frac{1}{2}}$; the energy density at the point of impact, taking into account the Fitzgerald contraction, and the change of frequency of the source, is

$$E = E_s\left[\frac{1+\frac{V}{c}}{1-\frac{V}{c}}\right], \tag{21}$$

where E_s is the energy density due to the point source in a stationary system of the same measured dimensions;[4] the length of the packet is reduced in the ratio $[1-(V/c)] : 1$. The energy content of the packet emitted in apparent time Δt is the product of these, or

$$E_s(c\Delta t)\frac{\left(1+\frac{V}{c}\right)}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}. \tag{22}$$

For the second of these changes we note that because of the Fitzgerald contraction and the (experimentally established) variation of clock rate with motion the absolute velocity which will, on the moving platform, be measured as v, is $(V+v)/[1+(vV/c^2)]$.

Our procedure is now to substitute the changed energy content in the equations of motion for a moving platform, and find the resultant velocity. We do this in the expression for the conservation of momentum, where the expression corresponding to (18) becomes

$$E_s\frac{(c\Delta t)}{c^2}\frac{\left(1+\frac{V}{c}\right)}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}c+\frac{M_0V}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}} = \frac{\left(M_0+\frac{H't'}{c^2}\right)v'}{\left[1-\frac{v'^2}{c^2}\right]^{\frac{1}{2}}}, \tag{23}$$

where v' is the velocity of the particle after the impact, and $H't'$ the heat absorbed by the particle. Now by the conservation of mass

$$\frac{M_0+\frac{H't'}{c^2}}{\left[1-\frac{v'^2}{c^2}\right]^{\frac{1}{2}}} = E_s\frac{(c\Delta t)}{c^2}\frac{\left(1+\frac{V}{c}\right)}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}+\frac{M_0}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}};$$

also, from (19)

$$M_0 = E_s\frac{(c\Delta t)}{cv}\left(1-\frac{v}{c}\right).$$

Making these substitutions in (23) we obtain

$$v' = \frac{\frac{\left(1+\frac{V}{c}\right)}{c}+\frac{\left(1-\frac{v}{c}\right)V}{cv}}{\frac{\left(1+\frac{V}{c}\right)}{c^2}+\frac{\left(1-\frac{v}{c}\right)}{cv}} = \frac{V+v}{1+\frac{vV}{c^2}}, \tag{24}$$

[4] Abraham, *Theorie der Elektrizität* (B. G. Teubner, Leipzig, 1905), Vol. 2, p. 383, Eq. (245).

showing that the *acquired velocity is that which will, on the moving platform, measure the same as that acquired on the stationary platform.*

To investigate the heat absorbed by the particle when the impact occurs on a moving platform, which we have designated $H't'$, we substitute in (23) the value we have just found for v', together with the value of M_0 from (19). From this operation we obtain

$$H't' = E_s(c\Delta t)\left[1-\frac{c}{v}\left(1-\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}\right)\right] \quad (25)$$

which is identical with (20), showing that *the heat absorbed by the particle is the same as for the case of the stationary platform.*

(3) PRESSURE OF RADIATION ON AN ABSORBING PARTICLE

Since the velocity acquired by an absorbing particle, and the heat absorbed, are the same for the cases of the stationary and moving platforms, it is reasonable to infer that the pressure of radiation is likewise the same in both cases. The matter should, however, be looked into more closely, since the commonly given values for radiation pressure on a moving absorber lead to a different conclusion.

We arrive at a straightforward solution for the pressure by assuming that the pressure is the quotient of the change of momentum of the incident energy by the time; as expressed in the equation, applying to the moving platform.

$$Pt = E_s\frac{(c\Delta t)}{c^2}\frac{\left(1+\frac{V}{c}\right)}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}c - \frac{Ht(V+v)}{c^2\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}. \quad (26)$$

Substituting the value

$$E_s(c\Delta t)\left\{1-\frac{c}{v}\left[1-\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}\right]\right\}$$

for $H't'$ this becomes

$$Pt = \frac{E_s(c\Delta t)}{c^2\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}\left[\left(1+\frac{V}{c}\right)c - \frac{(V+v)}{\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}\left\{1-\frac{c}{v}\left[1-\left(1-\frac{v^2}{c^2}\right)^{\frac{1}{2}}\right]\right\}\right]. \quad (27)$$

Proceeding at once to the case of co-moving source and target we have that $v\cong 0$, so that

$$Pt = \frac{E_s(c\Delta t)}{c\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}} \quad (28)$$

and

$$t\cong\frac{\Delta t}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}},$$

so that

$$P = E_s. \quad (29)$$

Remembering that the energy density E in this case is $E_s[1+(V/c)]/[1-(V/c)]$ we have the relation for the pressure on a moving absorbing surface

$$P = E\left[\frac{1-\frac{V}{c}}{1+\frac{V}{c}}\right], \quad (30)$$

which is to be contrasted with the value $P=E[1-(V/c)]$ uniformly given by previous investigators;[5] which, as pointed out in a previous paper, calls for a variation of radiation pressure with motion of the system.

We can arrive at this same value after an analysis of the ultimate fate of the incident energy. This energy does three things:

[5] This value for the pressure on a moving absorbing target was obtained by the writer in a previous paper [H. E. Ives, J. Opt. Soc. Am. **32**, 32 (1942), in which pertinent references are given] but without the detailed analysis provided by the present paper.

(1) It puts the constituent atoms of the particle into motion, showing as a rise in temperature of the particle.

(2) It imparts kinetic energy to the particle due to the increased mass of the moving atoms.

(3) It does work, the product of pressure times distance, on the original mass of the particle.

According to this analysis we can obtain the pressure from the change of momentum of the *original mass of the particle*, divided by the time of impact, or

$$P=\frac{\Delta\bar{M}}{t}=\frac{1}{t}\left\{\frac{M_0(V+v)}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}-\frac{M_0V}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}\right\}$$

$$=\frac{M_0\left\{v+V\left(1-\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}\right)\right\}}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}. \quad (31)$$

Expanding $[1-(v^2/c^2)]^{\frac{1}{2}}$, substituting the value of M_0 from (19), and proceeding to the case of the co-moving source and particle, when $t=\Delta t/[1-(v^2/c^2)]^{\frac{1}{2}}$, we get again,

$$P=E_s.$$

We have now all the factors to work out exactly the allocation of the energy incident on an absorbing particle or surface. The equation for the conservation of energy in the impact of a wave packet on a particle, the experiment being performed on a moving platform, is

$$E_s(c\Delta t)\frac{\left(1+\frac{V}{c}\right)}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}+\frac{M_0c^2}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}$$

$$=\frac{\left(M_0+\frac{H't'}{c^2}\right)c^2\left\{\frac{V+v}{1+\frac{Vv}{c^2}}\right\}}{\left[1-\frac{(V+v)^2}{c^2\left(1+\frac{vV}{c^2}\right)^2}\right]^{\frac{1}{2}}}. \quad (32)$$

Grouping terms in M_0 and $H't'$

$$E_s(c\Delta t)\frac{\left(1+\frac{V}{c}\right)}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}=M_0c^2\left\{\frac{1+\frac{vV}{c^2}-\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}\right\}$$

$$+H't'\left\{\frac{1+\frac{vV}{c^2}-\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}\left[1-\frac{v^2}{c^2}\right]^{\frac{1}{2}}}\right\}+H't'. \quad (33)$$

Substituting the value of M_0 from (19), and the value of $H't'$ from (25), expanding $[1-(v^2/c^2)]^{\frac{1}{2}}$, and proceeding to the case of $v\leqq 0$ we get

$$E_s(c\Delta t)\frac{\left(1+\frac{V}{c}\right)}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}} = \frac{E_s(c\Delta t)}{c\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}V + E_s(c\Delta t)\left[\frac{1}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}-1\right] + E_s(c\Delta t). \quad (34)$$

Energy in wave packet.	Pressure × time × velocity or mechanical work done by radiation on particle.	Kinetic energy added to particle by heat motion.	Heat motion of atoms, causing rise of temperature.

If we now divide through by the time $\Delta t/[1-(V^2/c^2)]^{\frac{1}{2}}$ we obtain the relation between the *rate* of flow of energy into the particle and the rate of distribution of this energy among its various tasks, as follows:

$$E_s(c+V)=PV+E_sc\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}\times\left\{\frac{1}{\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}}-1\right\}+E_sc\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}. \quad (35)$$

The last term indicates that the rate of heating (rise of temperature) is less than for the stationary platform by the factor $[1-(V^2/c^2)]^{\frac{1}{2}}$, but the reciprocal of this is exactly the factor by which the unit of time on the moving body is increased, so that the *rate* of heating will be observed the same. In short, the measurable quantities, pressure and rate of heating, appear invariant with the motion of the platform.

The last term of (35), giving the rate of heating, may be rewritten, in terms of V_s, the velocity of the source, and V_T the velocity of the target, as follows:

$$H=E_sc\left[\frac{1+\frac{V_s}{c}}{1-\frac{V_s}{c}}\right]\left[\frac{1-\frac{V_T}{c}}{1+\frac{V_T}{c}}\right]\left[1-\frac{V_T^2}{c^2}\right]^{\frac{1}{2}}. \quad (36)$$

From this the heating with either source or target alone moving can be derived. If the target is stationary this gives

$$H=E_sc\left[\frac{1+\frac{V}{c}}{1-\frac{V}{c}}\right],$$

or, if the source is moving *away* from the target

$$H=E_sc\left[\frac{1-\frac{V}{c}}{1+\frac{V}{c}}\right]. \quad (37)$$

If the target is moving away from the source

$$H=E_sc\left[\frac{1-\frac{V}{c}}{1+\frac{V}{c}}\right]\left[1-\frac{V^2}{c^2}\right]^{\frac{1}{2}}, \quad (38)$$

so that *if the rate of heating is measured on the target*, it is the same in each case.[6]

(4) DISCUSSION

In this study, and in the paper immediately preceding,[1] we have applied the assumptions of conservation of energy and momentum to the mass equivalent of radiation. These assumptions have led to the variation of mass with velocity in the ratio $1/[1-(v^2/c^2)]^{\frac{1}{2}}:1$, to values for the velocity and heat acquired by an absorbing particle under the impact of radiation, and for the pressure of radiation on absorbing targets. It is found that these assumptions lead to complete invariance of all measurable quantities (absorbed heat, rate of heating, velocity, pressure) with velocity of the system.[7]

It is to be recognized, however, that while the results obtained are self-consistent, the values of the various quantities, in particular of pressure on a moving absorbing target, are at significant variance with the values obtained by other approaches.

[6] There is also involved in this analysis (in evaluating the energy density from the source at the target) the convention that the distance between source and target is always measured at the target, on a scale attached to the source and experiencing the Fitzgerald contraction proper to the velocity of the source.

[7] Since, as pointed out by Epstein [Am. J. Phys. **10**, 1 (1942)] the Fitzgerald contraction has been derived from a simple law of attraction, and the variation of clock rate with velocity follows from the variation of mass with velocity, there appears to be no need of a principle of relativity to derive the general invariance of optical phenomena with motion.

The Spectrographic Determination of Small Amounts of Impurities in Magnesium Chloride*

MABEL F. WILSON
Spectrographic Department, Diamond Alkali Company, Painesville, Ohio

THE successful deposition of metallic magnesium requires a close control of the minor impurities in the magnesium chloride from which it is produced. Accurate chemical determinations of these impurities are difficult, time consuming, and, in some instances, of questionable accuracy. The spectrograph may be utilized to great advantage for the rapid and accurate analysis of such impurities as boron, silicon, iron, manganese, aluminum, copper, sodium chloride, and calcium chloride.

No chemical manipulation of the sample is necessary. The diluted sample containing the internal indicator is placed on purified graphite rods, dried, and exposures made by a.c. arc excitation. The plates are carefully processed, densitometered, and calculations made from analytical curves.

EQUIPMENT

A Bausch and Lomb large Littrow spectrograph with quartz optics, covering a wave-length range of 2100 to 8000 angstrom units is used. A spherical quartz condensing lens increases the intensity of the light from the arc, but is not sharply focused so that the images of the electrodes are not included. The variable slit is set at four microns.

An a.c. arc constructed in the laboratory, similar to the one described by Duffendack,[1] operates at 2300 volts, 2¼ amperes. A "Sola" voltage regulator on the primary circuit insures an even voltage output. The arc is encased in a safety cage, the circuit operating only when the door is closed. The lower electrode holder is fixed with respect to the mounting and the upper one operated by a screw adjustment. The entire cage is mounted on three adjustable legs for ease in leveling.

A prefocused projection lamp mounted so it may be moved readily into the optical path throws a sharp, enlarged image of the arc onto a white screen behind the cage, making the adjustment to 0.75 mm arc gap comparatively simple.

A Bausch and Lomb seven-step sector, ratio 1–1.5, operated by a special 2200-r.p.m. motor is mounted just in front of the slit.

A Hilger visible microphotometer is employed in the quantitative determinations of the samples. Galvanometer deflections are thrown on a 50-mm scale, which is adjusted so that a reading of zero is obtained when the slit is covered by the diaphragm and, by means of a rheostat, the galvanometer deflection is set at 50 when the clear plate is brought into focus. This instrument is operated by a storage battery, which is connected to a trickle charger except when readings are being made.

A home-made developing cabinet made from an electric refrigerator[2] maintains a constant developing temperature of 68°F plus or minus 0.5°. The stainless steel tanks containing the solutions are large enough to do away with the necessity of mechanical agitation during development. Eastman antihalation process plates are developed five minutes in D-19 developer and fixed in Formula F5.

PREPARATION OF MATERIALS

Redistilled metallic magnesium, 99.99 percent pure, was dissolved in hydrochloric acid to form a pure magnesium chloride solution for the preparation of standards. The specific gravity was adjusted to about 1.325 and the solution made alkaline to methyl red. A qualitative analysis under the conditions of exposure to be used showed no lines of the impurities for which determinations were to be made.

A water solution of ammonium molybdate was prepared as an internal indicator solution, so that each ml contained 2 mg of molybdenum. If no analysis is to be made for sodium chloride,

* Paper presented at the Twenty-Eighth Annual Meeting of the Optical Society of America, Pittsburgh, Pennsylvania, October 7–9, 1943.

[1] O. S. Duffendack and K. B. Thomson, Proc. A.S.T.M. 36, II, 301 (1936).

[2] F. Lee, Photo Tech., pp. 40–41 (June, 1941).

a sodium molybdate solution, containing the same concentration of molybdenum, may be used.

For each concentration of unknown prepared a 10-ml portion of the base solution and an equal volume of the internal indicator were pipetted into a 100-ml volumetric flask. The required amount of the impurities were added in the form of dilute solutions, the volume adjusted to give approximately 50 grams of magnesium chloride per liter, and the solution made just acid to methyl red with 1 *N* HCl. From 6 to 9 different

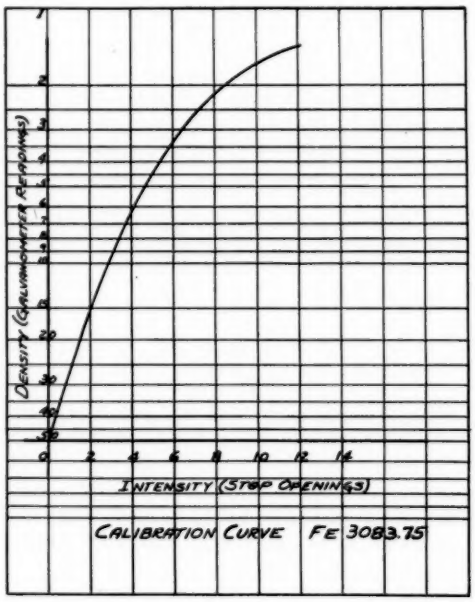


FIG. 1.

concentrations of each impurity were prepared, depending on the range to be covered. The amount of each impurity in the standard solutions was expressed in grams per liter.

The specific gravity of the routine samples is determined whenever possible by means of a Westphal balance. More concentrated solutions which are apt to crystallize when cold make necessary the use of a hydrometer at 60°C. A 10-ml portion of the clear solution is removed with a pipette, an equal amount of the internal indicator added, and the volume and acidity adjusted exactly as in the preparation of the standard solutions. For turbid solutions, the sample is increased to 100 ml, made acid to methyl red, diluted to 500 ml, and an aliquot equivalent to the original 10-ml sample is removed. Very dilute solutions are concentrated by evaporation before removing the sample. Solids are prepared by dissolving a weighed amount in water to give a solution approximately the same in specific gravity as the regular routine liquors.

National Carbon Company Special Spectroscopic graphite rods $\frac{1}{4}''$ by 12″ are cut in $\frac{1}{2}''$ lengths by a tool designed for the purpose. The thin, fine-toothed saw of this tool not only cuts the electrodes perpendicular to the axis of the rod but also polishes the cut surface. These lengths are placed in graphite holders, a pair for each exposure to be made, and each one coated with 0.05 ml of the sample, delivered from a 0.1-ml pipette. The coated electrodes are dried in an oven at 80–90°C for at least an hour before arcing. When dry, each has a smooth, thin, white coating well out to the edges.

EXPERIMENTAL PROCEDURE

The electrodes are placed in the holders one above the other and the upper one lowered until an arc is struck when the safety door is closed. A pre-burning time of 5–8 seconds serves to level off any uneven edges on the electrodes. After this pre-burning period, the arc gap is adjusted to exactly 0.75 mm and this width maintained throughout the exposure. Only one exposure is made on the same pair of electrodes, but each sample is exposed in triplicate. Exposures of thirty seconds in the wave-length range 2650 to 8000 angstrom units, and three minutes in the range 2350 to 3200, are sufficient to cover all the impurities in the range desired.

An iron arc, operated on 220 volts $2\frac{1}{2}$ amperes d.c., is exposed for ninety seconds through the step sector without an external lens. The iron line 3083.75A covers a satisfactory density range on a process plate and may be used for all the lines measured as the H and D curves in the region 3100 to 2500A are parallel.[3] The completed calibration curve is shown in Fig. 1. Calibration of each plate is not necessary as constant develop-

[3] W. C. Pierce and N. H. Nachtrieb, Ind. Eng. Chem. Anal. Ed. **13**, 774–781 (1941).

ing conditions are maintained. Recalibration of each new emulsion number is sufficient.

With the scale adjusted properly for zero and 50, galvanometer readings are made on the impurity and internal indicator lines selected for analysis. Background readings are necessary only on the three-minute exposures.[3,4] Proper location for each background reading was determined experimentally and must be taken in exactly the same position each time.

The selection of line pairs for analysis was carried out as nearly as possible to meet the requirements of such pairs as mentioned by Pierce and Nachtrieb,[3] Twyman,[5] Gerlach and Schweitzer,[6] and others. In brief, these requirements are: Both lines should arise from the same method of excitation and be of the same order; an equality of intensity should occur near the middle of the concentration range; the lines should be free of interference by extraneous elements and from excessive backgrounds; they should be in the same region of the plate. In addition to meeting these requirements, the results on repeated runs of the same sample were checked experimentally and the pairs giving the best actual duplication were selected.

Preliminary analysis of samples from a number of sources determined the concentrations to be covered. Table I shows the range of impurities covered by the analytical curves calculated on the line pairs shown in the second column.

The calibration curve is plotted by a method similar to that described by Pierce and Nachtrieb, in the article previously mentioned, for use when background corrections are necessary. Logarithms of the galvanometer deflections are plotted as ordinates, against the relative intensities of exposure, represented by step opening ratios, as abscissae on a linear scale. Background corrections are made by subtraction of background intensity from line intensity. The corrected line intensity of the impurity is divided by that of the internal indicator to obtain the intensity ratio for each concentration of the standard solutions. The log of this ratio is plotted against the log of concentration (in grams per liter) to obtain the analytical curve. Figure 2 shows this curve for boron. It is a straight line within the limits of concentration shown in the figure. Table II illustrates how the calculation of the amount of boron in a sample is determined from the curve. Calculations of other impurities are obtained in the same manner.

TABLE I.

Impurity	Line pair	Exposure in seconds	g/l impurity
NaCl	Na 3302.4A / Mo 3208.9A	30	0.5 −30.0
$CaCl_2$	Ca 3181.3A / Mo 3208.9A	30	0.5 −30.0
	Ca 3158.9A / Mo 3208.9A	30	0.25 −10.0
Cu	Cu 3247.5A / Mo 3208.9A	30	0.0008−0.2
	Cu 3273.9A / Mo 3208.9A	30	0.002 −0.5
Al	Al 3082.2A / Mo 3116.1A	90	0.02 −2.0
B	B 2497.7A / Mo 2538.5A	90	0.0002−0.4
Si	Si 2506.9A / Mo 2538.5A	90	0.001 −0.2
	Si 2516.2A / Mo 2538.5A	90	0.005 −0.5
Fe	Fe 2599.4A / Mo 2538.5A	90	0.0008−0.2
	Fe 2563.5A / Mo 2538.5A	90	0.02 −5.0
Mn	Mn 2605.7A / Mo 2538.5A	90	0.0002−0.03

DISCUSSION OF RESULTS

The final test of the efficiency of any method is, of course, in its everyday use. For the past

[4] J. Cholak and R. V. Story, J. Opt. Soc. Am. **32**, 502–505 (1942).
[5] F. Twyman, J. Inst. Metals **64**, Advance Copy No. 828 (1939).
[6] W. Gerlach and E. Schweitzer, *Foundations and Methods of Chemical Analysis by the Emission Spectrum* (Adam Hilger, London, 1931), p. 67.

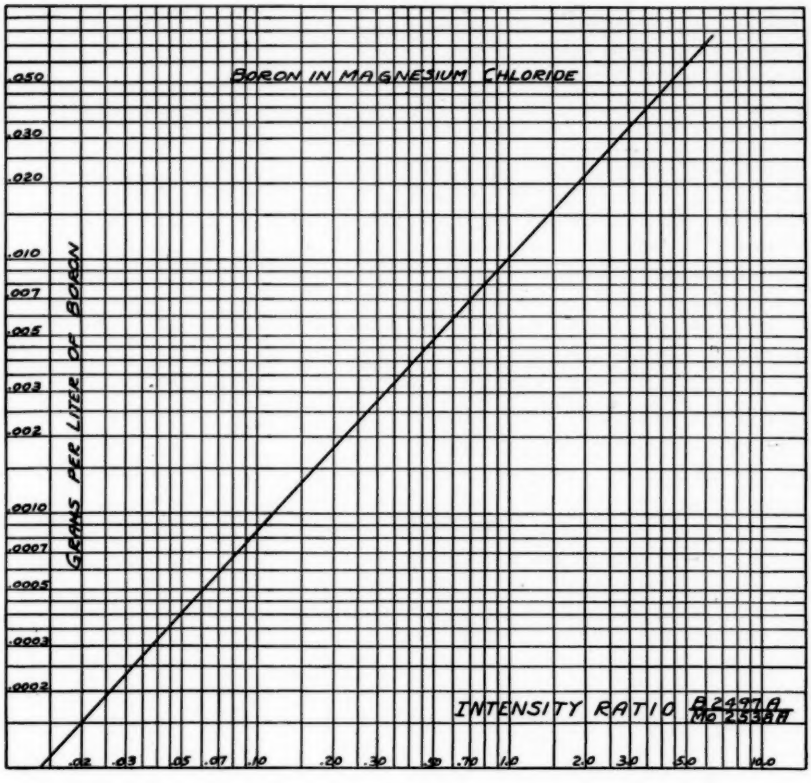


FIG. 2.

year, daily analyses of the routine magnesium chloride liquor samples have been carried out with gratifying results. An accuracy of 10 percent of the amount present has been maintained, four different operators making the determinations with no detectable variation in results.

The time required for the complete analysis of a sample in triplicate is 2½ hours, an average of 5½ minutes per determination. This time can be cut to 4 minutes when two samples are run on the same plate.

Table III gives some typical results obtained on one of the impurities mentioned: $CaCl_2$. The figures were taken from the regular daily records of two different chemical laboratories and our spectrographic laboratory. The chemical results were obtained by two different methods. The three groups, selected to show the equal accuracy in different ranges of concentration, are not continuous; but the figures in each group are consecutive. Those in boldface italic type have been omitted from the averages shown in the right-hand column.

TABLE II. Calculation of results.

B g	Line	B g	Line
47.0	25.5	44.3	9.8
45.5	22.8	44.5	8.8
43.0	21.0	41.2	7.7
B	47.0−0.05	Difference	−0.85
	25.5−0.90		
Mo	44.3−0.20	Difference	−2.5
	9.8−2.7		

B/Mo = 0.85/2.5 = 0.34 intensity ratio

Similarly, other ratios are 0.35 and 0.36, giving an average of 0.35 for the three determinations. 0.35 corresponds to 0.0033 g/l of B on the analytical curve. 0.0033/1170.60 = 0.000274 percent B, or 2.74 p.p.m. B in that sample.

The same degree of accuracy has been maintained for samples of low concentrations and for solids, as for the routine liquors of more consistent concentration.

SUMMARY

In conclusion, the data as given in this paper show that:

(1) Analytical curves have been worked out for boron, silicon, manganese, iron, aluminum, copper, sodium chloride, and calcium chloride.

(2) Samples of magnesium chloride varying from very dilute solutions to solids may be analyzed on these same curves with the same high degree of accuracy.

(3) The time required per determination recommends the method for routine use.

(4) The method compares favorably with chemical analysis in accuracy and reproducibility.

(5) Satisfactory results have been obtained in daily routine use, over a period of more than a year.

TABLE III. Percent $CaCl_2$ in $MgCl_2$.

Chem. No. 1	Chem. No. 2	Spectro.	Average
1.00	0.96	0.98	0.98
0.98	0.93	0.97	0.96
1.04	1.02	1.00	1.02
1.10	1.07	1.14	1.10
0.84	0.78	0.81	0.81
0.86	0.79	0.80	0.82
0.25	0.20	0.24	0.23
0.24	0.20	0.23	0.22
0.15	***0.08***	0.17	0.16
0.31	0.24	0.29	0.28
0.28	***0.18***	0.26	0.27
0.43	***0.35***	0.45	0.44
0.32	0.28	0.35	0.32
0.22	0.17	0.22	0.21
0.51	0.53	0.50	0.51
0.46	0.54	0.50	0.50
0.41	0.51	0.47	0.45
0.42	***0.62***	0.44	0.43
0.48	0.54	0.45	0.49
0.55	0.64	0.66	0.65
0.46	0.64	0.60	0.62

Aesthetic Measure Applied to Color Harmony*

PARRY MOON, *Massachusetts Institute of Technology, Cambridge, Massachusetts*

AND

DOMINA EBERLE SPENCER, *Tufts College, Medford, Massachusetts*

1. INTRODUCTION

THE problem of why some things are beautiful and others are not has been the subject of much speculation. Plato said,[1] "Measure and proportion always pass into beauty and excellence." The Greek view[2] was that "beauty consists in the . . . expression of *unity in variety*." This basic idea is still alive today[3] and is perhaps the most firmly established principle of modern aesthetics.

While it is true that many works of art appeal largely because of their connotative associations, it is also true that much beauty is of a more formal nature, as exemplified by regular geometric shapes, by musical chords, and by related colors. There is hope that this latter type of beauty is simple enough to be analyzed, and it is to this subject that we apply the idea that beauty is "order in complexity." Fechner[4] stated the principle in 1876, but it was first placed on a quantitative basis by G. D. Birkhoff[5] in his *theory of aesthetic measure* (1933).

Birkhoff applied his theory to ornaments, vases, music, and poetry; but he did not apply it to color. The purpose of the present paper is to show that aesthetic measure can be employed to good advantage in color harmony.

2. AESTHETIC MEASURE

According to Birkhoff, aesthetic measure is defined by the equation,

$$M = O/C, \tag{1}$$

where O represents the number of elements of order and C is the number of elements of complexity. Evidently a high aesthetic measure can be obtained with a very simple design (small C), though still higher measures may be possible with a complex design having a sufficiently orderly arrangement (large O). Equation (1) is a mathematical statement of the principle given in the preceding section. Other equations would express the principle equally well, and only a great amount of further experimental research can settle the question of whether the simple form of Eq. (1) is best.

The problem is to determine O and C for each class of aesthetic objects. With polygonal forms, for example, Birkhoff defined the complexity C as equal to the number of indefinitely extended lines that contain all the sides of the polygon. The elements of order are vertical symmetry, equilibrium, rotational symmetry, and relation to a horizontal-vertical coordinate system. To each of these elements is assigned a weight of 0, 1, or 2. If the form is unsatisfactory in any respect, 1 or 2 is subtracted from the sum of the previous elements of order. The choice of these particular numerical factors was apparently determined by a process of trial and error, conditioned by a desire to maintain the greatest possible simplicity.

Ninety polygons were evaluated by Birkhoff. The aesthetic measures, calculated from Eq. (1), ranged from 1.50 for a square to −0.17 for an

---

* Paper presented at the Twenty-Eighth Annual Meeting of the Optical Society of America, Pittsburgh, Pennsylvania, October 7–9, 1943.

[1] Plato, *Philebus*, in *Dialogues of Plato*, Jowett translation.

[2] B. Bosanquet, *A History of Aesthetic* (London, 1892).

[3] W. Abell, *Representation and Form* (Scribner's Sons, New York, 1936); E. S. Baker, Univ. Toronto Studies, Psychol. Series, No. 4, p. 205 (1900); Cattel, Glascock, and Washburn, Am. J. Psychol. **29**, 333 (1918); B. Croce, *Aesthetic as Science of Expression* (London, 1922); H. Davies, Phil. Rev. **10**, 28 (1901); J. Dewey, *Art as Experience* (New York, 1934); J. Evans, *Taste and Temperament* (London, 1939); K. E. Gilbert and H. Kuhn, *A History of Aesthetics* (Macmillan, New York, 1939); K. Gordon, *Aesthetics* (Henry Holt, New York, 1909); A. Kirschmann, Univ. Toronto Studies, Psychol. Series, p. 177 (1900); C. Lalo, *L'Esthétique Expérimentale Contemporaine* (Paris, 1908); H. S. Langfeld, *The Aesthetic Attitude* (Harcourt Brace, New York, 1920); W. Lietzmann, *Mathematik und Bildende Kunst* (Breslau, 1931); T. Munro, *Scientific Method in Aesthetics* (W. W. Norton, New York, 1928); K. E. Neuhaus, *The Appreciation of Art* (Ginn and Company, Boston, 1924); D. W. Ross, *A Theory of Pure Design* (Boston, 1907); G. Santayana, *The Sense of Beauty* (New York, 1896).

[4] G. T. Fechner, *Vorschule der Aesthetik* (Leipzig, 1876).

[5] G. D. Birkhoff, *Aesthetic Measure* (Harvard University Press, Cambridge, Massachusetts, 1933).

irregular six-sided figure. R. C. Davis[6] attempted an experimental check in 1936. The order of choice of the polygons, as determined by a group of observers, was not at all like that predicted by Eq. (1) with the constants of Birkhoff. This lack of agreement, however, does not prove that Birkhoff's method is wrong but may indicate that his numerical weights should be revised. In fact, Rashevsky[7] has used a modification which gives fairly good agreement with the experimental data of Davis. A number of other investigators have subjected Birkhoff's classification of polygons to experimental test.[8]

Birkhoff applied Eq. (1) to polygonal shapes, ornaments, shapes of vases, musical harmony and melody, and poetry. Of color he says, "The simpler the palette is, the less will be the complexity, so that the palette should be as restricted as the subject permits. Evidently the eye appreciates the repetition of a color, a graded sequence of colors, and a balance of light and dark values about the centers of interest, in accordance with the classification of Ross. All of these elements of order are of definite aesthetic importance." Beyond this brief mention of color[9] there seems to have been no previous application of aesthetic measure to color harmony.

The present paper applies aesthetic measure to give a rating system for color harmonies. Elements of order are identity, similarity, and contrast of hue, value, and chroma. Less desirable relations are the ambiguities treated in a previous paper.[10] Area balance[11] also constitutes an element of order. The numerical weights for the various elements are determined by experiment, and the resulting system allows color harmonies to be developed and compared theoretically without actually producing them.

3. IDENTITY, SIMILARITY, AND CONTRAST

In applying aesthetic measure to color harmony, one must decide on specific methods of evaluating order and complexity. Birkhoff, in his various applications, seems to have used little experimental work and to have assigned weights of 0, ± 1, ± 2 in a rather arbitrary manner. An attempt to use these weights in color harmony showed the procedure to be insufficiently refined. It was necessary to determine the various factors experimentally. A method of evaluating complexity C must be established and the aesthetic importance of identity, similarity, and contrast must be determined. The quantitative effect of ambiguities and area balance must then be included.

The simplest definition of the complexity C is obtained by equating it to the number of colors in the design. Thus for a 2-color design, $C=2$ and for an arrangement of three colors, $C=3$. But this definition was unsatisfactory, for with it no simple expression that would fit the experimental results could be found for the measure M. A satisfactory definition of complexity is

$$C=(\text{No. of colors})+(\text{No. pairs of colors having hue difference})+(\text{No. pairs with value difference})+(\text{No. pairs with chroma difference}). \quad (2)$$

For example, with the pair of colors $R\,5/6$, $R\,4/2$, the number of colors is 2, the number of pairs having hue difference is 0, the number of pairs with value difference is 1, and the number of pairs with chroma difference is 1. Thus,

$$C=2+0+1+1=4.$$

Or consider the 3-element harmony $P\,3/2$, $P\,4/2$, $Y\,8/10$, where hue, value, and chroma are all variable. The number of colors is 3, the number of pairs having hue difference is 2, value difference gives 3, while chroma difference gives 2. Therefore, $C=10$.

The next factor to be evaluated is the order O, which must be determined for identity, similarity, and contrast in all three variables. A set of samples was made up to illustrate variation in hue only, in value only, in chroma only, and in combinations of these variables. The predominant hue was the same for all the samples of the set. Other sets were employed, each with a different predominant hue.

In the first test, each sample consisted of two color chips which were of equal area and which

[6] R. C. Davis, J. Gen. Psychol. **15**, 231 (1936).
[7] N. Rashevsky, *Advances and Applications of Mathematical Biology* (University of Chicago Press, 1940), p. 190.
[8] Beebe-Center and Pratt, J. Gen. Psychol. **7**, 339 (1937); Harsh, Beebe-Center, and Beebe-Center, J. Psychol. **7**, 343 (1939); H. J. Eysenck, Psychol. Rev. **48**, 83 (1941).
[9] Reference 5, p. 213.
[10] P. Moon and D. E. Spencer, J. Opt. Soc. Am. **34**, (1944).
[11] P. Moon and D. E. Spencer, J. Opt. Soc. Am. **34**, (1944).

TABLE I. Samples for identity and similarity.

Sample No.	Designation	Munsell samples R	Y	G	P
1	*HC, V*	*R* 7/8, 6/8	*Y* 8/8, 7/8	*G* 6/4, 5/4	*P* 5/8, 4/8
2	*G, V*	*N* 7, 6	*N* 8, 7	*N* 7, 6	*N* 4, 3
3	*VC, H*	*R* 7/8, *YR* 7/8	*Y* 8/8, *GY* 8/8	*G* 6/4, *BG* 6/4	*P* 5/8, *RP* 5/8
4	*HV, C*	*R* 7/4, 7/8	*Y* 8/8, 8/12	*G* 5/4, 5/8	*P* 5/8, 5/12
5	*H, VC*	*R* 7/6, 6/4	*Y* 8/6, 7/4	*G* 6/4, 5/2	*P* 5/6, 4/4
7	*V, HC*	*R* 7/8, *YR* 7/4	*Y* 8/8, *GY* 8/4	*G* 6/6, *BG* 6/2	*P* 5/8, *RP* 5/4
8	*C, HV*	*R* 6/6, *YR* 7/6	*Y* 7/6, *GY* 8/6	*G* 5/6, *BG* 6/6	*P* 4/6, *RP* 5/6
9	*—HVC*	*R* 7/6, *YR* 6/4	*Y* 8/6, *GY* 7/4	*G* 6/4, *BG* 5/2	*P* 5/6, *RP* 4/4
11	*HVC—*	*R* 6/8, 6/8	*Y* 8/8, 8/8	*G* 6/4, 6/4	*P* 5/8, 5/8

were chosen to avoid ambiguity.[10] Samples were viewed on a neutral background of Munsell value 5. Each pair of chips was mounted on a separate card, and the observer arranged the cards in what appeared to him to be the order of aesthetic merit. Several observers took part in the tests, including a group of architectural students who had had considerable training in color.

The first tests determined the factors for *identity* and *similarity*. The samples are specified in Table I. Eleven samples in each set were employed at first, but it was found that more reproducible results could be obtained with nine samples per set. An abbreviated designation is used in Table I; the letters before the comma refer to the quantities that are identical, while the letters after the comma refer to quantities that are varied by a small step.[10] Samples numbered 2 are achromatic, while the others are chromatic.

The average of experimental results gave the following order of preference for the most important samples:

No. 1—*HC, V;*
No. 5—*H, VC;*
No. 11—*HVC—*;
No. 4—*HV, C;*
No. 3—*VC, H;*
No. 7—*V, HC.*

These results give the order of aesthetic merit for the samples. The absolute values on the aesthetic measure scale, however, are not fixed but are somewhat arbitrary. Suppose that a scale is fixed by taking the following numerical designations for M:

Sample No. 1, $M=1$;
Sample No. 11, $M=0.5$;
Sample No. 3, $M=0.2$;
Sample No. 7, $M=-0.025$.

These designations are in the experimentally determined order, and their spacing corresponds roughly to the apparent aesthetic difference in the samples.

Table II lists the foregoing aesthetic measures and the calculated values of complexity and order. Let the desired factors for identity be called H_1, V_1, C_1 and the factors for similarity be H_2, V_2, C_2. Then the figures in the right-hand column of Table II lead to the following equations:

$$H_1+V_1+C_1=1.0, \tag{3}$$

$$H_1+C_1+V_2=3.0, \tag{4}$$

$$V_1+C_2+H_2=-0.10, \tag{5}$$

$$C_1+V_1+H_2=0.60. \tag{6}$$

These equations give the relations,

$$\begin{aligned} H_2&=H_1-0.4,\\ V_2&=V_1+2,\\ C_2&=C_1-0.7. \end{aligned} \tag{7}$$

Values of H_1, V_1, and C_1 are now selected, such that Eq. (3) is satisfied. These selected numbers are then substituted in Eq. (7) to give the remaining three constants. Various selections were tried, and final results which satisfy the experiments of Table I for all chromatic samples are

$$\begin{aligned} H_1&=1.5, & H_2&=1.1,\\ V_1&=-1.3, & V_2&=0.7,\\ C_1&=0.8; & C_2&=0.1. \end{aligned}$$

These constants determine the aesthetic measures of all the samples of Table I except the achromatic ones. For harmonies of grays, it is necessary to obtain an additional number, which is taken as $G_1=1$.

TABLE II. Postulated aesthetic measures for identity and similarity.

Sample No.	Designation	C	M	$O=MC$
1	*HC, V*	3	1.0	3.0
11	*HVC—*	2	0.5	1.0
3	*VC, H*	3	0.2	0.6
7	*V, HC*	4	−0.025	−0.1

It is now possible to calculate the aesthetic measure of any harmony exhibiting identity and similarity. For example, No. 5 of Table I has identity of hue and small steps in value and

TABLE III. Calculated aesthetic measures for identity and similarity.

Sample No.	Designation	C	O	M
1	*HC, V*	3	3.0	1.00
8	*C, HV*	4	2.6	0.65
5	*H, VC*	4	2.3	0.58
2	*G, V*	3	1.7	0.57
11	*HVC−*	2	1.0	0.50
9	*−HVC*	5	1.9	0.38
3	*VC, H*	3	0.6	0.20
4	*HV, C*	3	0.3	0.10
7	*V, HC*	4	−0.1	−0.025

TABLE IV. Samples illustrating contrast.

Sample No.	Designation	Munsell samples *GY*	*PB*	*YB*
1	*HC, V*	*GY* 7/4, 3/4	*PB* 7/4, 3/4	*YR* 7/4, 3/4
2	*G, V*	*N* 7, 3	*N* 7, 3	*N* 7, 3
3	*VC, H*	*GY* 7/4, *P* 7/4	*PB* 7/4, *Y* 7/4	*YR* 7/4, *B* 7/4
4	*HV, C*	*GY* 7/2, 7/10	*PB* 5/2, 5/10	*YR* 6/2, 6/10
5	*H, VC*	*GY* 7/8, 4/2	*PB* 7/4, 3/12	*YR* 7/8, 4/2
7	*V, HC*	*GY* 7/10, *P* 7/2	*PB* 7/2, *Y* 7/10	*YR* 7/10, *B* 7/2
8	*C, HV*	*GY* 3/2, *P* 7/2	*PB* 3/2, *Y* 7/2	*YR* 3/2, *B* 7/2
9	*−HVC*	*GY* 7/8, *P* 4/2	*PB* 4/2, *Y* 7/8	*YR* 7/8, *B* 4/2
11	*HVC−*	*GY* 7/6, 7/6	*PB* 7/6, 7/6	*YR* 7/4, 7/4

chroma. The order O is equal to $H_1+V_2+C_2 = 1.5+0.7+0.1 = 2.3$. The complexity C is 4. So the aesthetic measure is

$$M = 2.3/4 = 0.58.$$

The measures of all the samples are listed in Table III. Evidently M can be either positive or negative. However, good harmonies have positive aesthetic measures, generally greater than 0.5, and only very bad color combinations have negative aesthetic measures. Thus the range of M for color harmonies is similar to the range used by Birkhoff for polygonal shapes.

One may wonder at the *negative* value found for V_1. Experiment shows that this negative factor is necessary for identity of value. Though identity of hue or of chroma is pleasing and results in a positive element of order, identity of value is found to be much worse than either similarity or contrast of value.

Identity and similarity have been evaluated, and the next step is the consideration of *contrast*. Table IV lists the samples used in the experiments with large steps. The order of preference was found to be essentially the same as with the previous samples exhibiting small steps. The set of equations for contrast was taken as

$$\begin{aligned} H_1+V_1+C_1 &= 1.0, \\ H_1+C_1+V_3 &= 6.0, \\ V_1+H_3+C_3 &= 0.8, \\ C_1+V_1+H_3 &= 1.2, \end{aligned} \tag{8}$$

from which,

$$\begin{aligned} H_3 &= H_1+0.2, \\ V_3 &= V_1+5.0, \\ C_3 &= C_1-0.4. \end{aligned} \tag{9}$$

The constants for contrast (Table V) were made larger than for similarity because experiment indicated that contrast was generally more pleasing than similarity. Since the quantities with subscripts 1 in Eq. (9) are already known, the new factors are found to be

$$\begin{aligned} H_3 &= 1.7, \\ V_3 &= 3.7, \\ C_3 &= 0.4. \end{aligned} \tag{10}$$

Calculated aesthetic measures for contrasting samples are listed in Table VI. Obviously the order of merit will not be exactly the same as for the similarity samples (Table III) since, for instance, G_1 is the same as before while $H_3 \neq H_2$.

TABLE V. Postulated aesthetic measures for contrast.

Sample No.	Designation	C	M	$O=MC$
1	*HC, V*	3	2.0	6.0
11	*HVC−*	2	0.5	1.0
3	*VC, H*	3	0.4	1.2
7	*V, HC*	4	0.2	0.8

TABLE VI. Calculated aesthetic measures for contrast.

Sample No.	Designation	C	O	M
1	*HC, V*	3	6.0	2.00
2	*G, V*	3	4.7	1.57
8	*C, HV*	4	6.2	1.55
5	*H, VC*	4	5.6	1.40
9	*−HVC*	5	5.8	1.16
11	*HVC−*	2	1.0	0.50
3	*VC, H*	3	1.2	0.40
4	*HV, C*	3	0.6	0.20
7	*V, HC*	4	0.8	0.20

Nevertheless, the theoretical order (Table VI) appears to be in accord with the average order in which the samples are arranged by competent observers.

TABLE VII. Calculated aesthetic measures, three-color harmonies.

Sample No.	Designation	Typical colors	C	O	M
1	*HC, V*	*Y* 8/8, 7/8, 6/8	6	8.1	1.35
5	*H, VC*	*Y* 8/8, 7/6, 6/4	9	5.9	0.66
9	*−HVC*	*Y* 8/8, *GY* 6/4, *YR* 7/6	12	4.25	0.35
3	*VC, H*	*Y* 8/8, *GY* 8/8, *YR* 8/8	6	1.35	0.23
4	*HV, C*	*Y* 8/4, 8/8, 8/12	6	1.2	0.20
7	*V, HC*	*Y* 8/8, *GY* 8/4, *YR* 8/4	8	−0.05	−0.01

TABLE VIII. Calculated aesthetic measures, two-color harmony with ambiguity.

1st Ambiguity in	Constant	C	O	M
H	*V, C*	3	−0.5	−0.16
V	*H, C*	3	+1.3	+0.43
C	*H, V*	3	+0.2	+0.07
HVC	—	5	−1.0	−0.20
2nd Ambiguity in				
H	*V, C*	3	+0.15	+0.05
V	*H, C*	3	+2.1	+0.70
C	*H, V*	3	+0.20	+0.07
HVC	—	5	+0.45	+0.09

A few experiments were also made with three-color harmonies. Aesthetic measure was calculated by the foregoing method and was found to be in satisfactory agreement with experiment. The calculated results are given in Table VII. For example, with three colors having the same hue but differing in value and chroma by small steps, the complexity is $3+0+3+3=9$. The order O is

$$3\,H_1+3\,C_2+3\,V_2=6.9.$$

Therefore, $M=6.9/9=0.76$. In some cases, ambiguities occur in the step between the first and third colors, even though the intervals between the first and second colors and between the second and third are legitimate small steps. The effects of these ambiguities have been included in the results of Table VII by a method that is described in the next section.

4. AMBIGUITIES

In the first paper of this series,[10] one of the postulates dealt with ambiguities. This subject will now be considered as a part of the theory of aesthetic measure. Samples were made to illustrate the various ambiguities of hue, value, and chroma. These samples, together with the previous ones, were arranged in order of preference, and in this way the effect of ambiguities was evaluated. For first ambiguities—that is, ambiguities between identity and similarity—the factors are called H_{A1}, V_{A1}, C_{A1}. For second ambiguities, the corresponding factors are called H_{A2}, V_{A2}, C_{A2}. The numerical results obtained from experiments are:

$$H_{A1}=0, \quad V_{A1}=-1.0, \quad C_{A1}=0;$$
$$H_{A2}=0.65, \quad V_{A2}=-0.20, \quad C_{A2}=0.$$

These results were used in calculating some of the aesthetic measures of Table VII. Table VIII lists the calculated aesthetic measures used in obtaining the foregoing constants. The third sample, for instance, is similar to Sample No. 1 of Table III, but the legitimate small step has been replaced by a first ambiguity. The fourth sample has ambiguities in H, V, and C. The complexity is therefore $C=2+3=5$. The order is $O=H_{A1}+V_{A1}+C_{A1}=-1.0$, and the aesthetic measure is $M=-1.0/5=-0.2$.

5. AREA BALANCE

The subject of area balance was considered in a previous paper.[11] Numerical factors must now be evaluated to indicate the additional order obtained when the areas are correctly adjusted. A pleasing effect is found[11] when the scalar moment about the adaptation point in ω space is equal for each color. Other pleasing effects are obtained when the moments are simple multiples of each other.

Ordinarily, equal moments give the highest aesthetic effect, while lower rating should be given to ratios of 2 or 3. It is found that satisfactory weights for these elements of order are

$$A_1=1.0, \quad A_2=0.50, \quad A_3=0.25$$

where the subscripts denote the ratio of scalar moments.

For example, consider a harmony of contrast, No. 3 of Table V. Without area balance, it has an aesthetic measure of 0.4. The complexity is 3 and the order is 1.2. With areas adjusted to give equal moments for the two color patches, an additional element of order is introduced and

$$O=1.2+1.0=2.2,$$
$$M=2.2/3=0.73.$$

Other calculated aesthetic measures are listed in Tables IX and X.

This completes the description of the method used in the experimental evaluation of aesthetic-measure factors. In the past, the comparison of art objects has been done by two methods—the method of the artist, who tries and compares and uses his own judgment; and the method of the experimental psychologist, who applies statistical analysis to the findings of a group of observers. We have chosen a middle ground, with some formal experiments combined with a large number of trials and comparisons of which no record was kept. The final scheme has been tested by calculating aesthetic measures for some fifty colored designs and comparing these values of M with direct estimates. In all cases the new method seems to give a valid aesthetic rating. There is much room for further improvement, however, particularly by the use of a large group of observers and a more refined experimental technique.

TABLE IX. Calculated aesthetic measures for area balance and similarity.

Sample No.	Designation	C	Aesthetic measure M Without balance	1 : 1	1 : 2	1 : 3
1	*HC, V*	3	1.00	1.33	1.17	1.08
8	*C, HV*	4	0.65	0.90	0.78	0.71
5	*H, VC*	4	0.58	0.83	0.70	0.64
2	*G, V*	3	0.57	0.90	0.73	0.65
11	*HVC−*	2	0.50	—	—	—
9	*−HVC*	5	0.38	0.58	0.48	0.43
3	*VC, H*	3	0.20	0.53	0.37	0.28
4	*HV, C*	3	0.10	0.43	0.27	0.18
7	*V, HC*	4	−0.025	0.23	0.10	0.04

TABLE X. Calculated aesthetic measures for area balance and contrast.

Sample No.	Designation	C	Aesthetic measure M Without balance	1 : 1	1 : 2	1 : 3
1	*HC, V*	3	2.00	2.33	2.17	2.08
2	*G, V*	3	1.57	1.90	1.73	1.65
8	*C, HV*	4	1.55	1.80	1.68	1.61
5	*H, VC*	4	1.40	1.65	1.53	1.46
9	−HVC	5	1.16	1.36	1.26	1.21
11	*HVC−*	2	0.50	—	—	—
3	*VC, H*	3	0.40	0.73	0.57	0.48
4	*HV, C*	3	0.20	0.53	0.37	0.28
7	*V, HC*	4	0.20	0.45	0.33	0.26

6. METHOD OF CALCULATION

The results of the preceding sections are summarized in Table XI, which gives all the numerical values that are necessary for the calculation of aesthetic measure. The procedure used in the calculations will now be described. The complexity C is first determined from Eq. (2), the effect of area balance is considered next, and then the elements of order associated with hue, value, and chroma are evaluated. Finally the aesthetic measure M is calculated by use of Eq. (1).

TABLE XI. Aesthetic factors.

	Identity	1st ambiguity	Similarity	2nd ambiguity	Contrast	Glare
Hue	+1.5	0	+1.1	+0.65	+1.7	—
Value	−1.3	−1.0	0.7	−0.20	3.7	−2.0
Chroma	+0.8	0	0.1	0	0.4	—
Gray	+1.0					

Area balance: Identity of moments, +1.0; moment ratio of ½ or 2, 0.5; moment ratio of ⅓ or 3, 0.25; other ratios, 0.00.

TABLE XII.

	Identity	1st ambiguity	Similarity	2nd ambiguity	Contrast	Glare
Hue						
Value						
Chroma						

The effect A of area balance is obtained by determining the scalar moment of each color patch.[11] Equality of moments of a pair of color patches contributes +1.0 to the order O, while moment ratios of 2 and 3 contribute 0.5 and 0.25 (Table XI). The total effect A of area is obtained by adding these factors for all possible pairs of colors. In a 2-color design with equal moments about the adaptation point, the result is $A=+1.0$. For a 3-color design there are 3 pairs; and therefore if the moments are equal, $A=3.0$. For a 4-color design, the result is 6; and for n colors it is equal to the number of combinations of n things taken 2 at a time, or

$$A=C^n{}_2=n!/[2!(n-2)!]. \qquad (11)$$

If some of the moment ratios are not unity, the appropriate factors of Table XI must be used.

Now calculate the elements of order associated with hue, value, and chroma. The easiest way of making the computation is to prepare a table (Table XII). All the pairs of colors are then considered, one pair at a time, and check marks are placed at the appropriate places in the table. For hue, the correct classifications are found by use

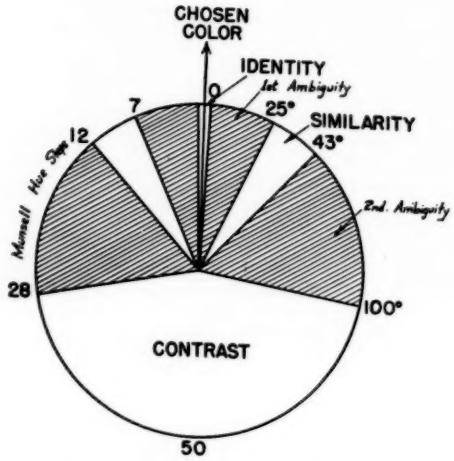


FIG. 1a. Regions of ambiguity, similarity, and contrast in hue.

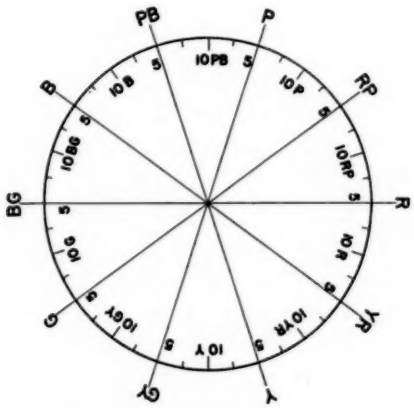


FIG. 1b. Munsell hue circle (to be used with Fig. 1a).

of Fig. 1. For a variation in chroma only or value only, the classification is obtained from Table I or II of the first paper of this series.[10] With simultaneous variation in value and chroma, Fig. 4 of the first paper[10] is necessary, and check marks are placed in the same column for both value and chroma. The weights of all elements of order are then obtained from Table XI.

The method is understood most easily by reference to a numerical example. Take a 2-color design with *R* 6/8, *R* 5/8 and with areas adjusted to give equal scalar moments about the adaptation point (*N* 5). Then

Number of colors	=2
Pairs of colors differing in hue	=0
Pairs of colors differing in value	=1
Pairs of colors differing in chroma	=0
	3.

Thus, according to Eq. (2), the complexity C is 3.

Area balance gives an additional element of order, $A=1$. Consideration of hue, value, and chroma gives Table XIII. Thus, from Table XI, the order is

$$O=A+1(1.5)+1(0.7)+1(0.8)=4.0$$

and the aesthetic measure is

$$M=O/C=4.0/3=1.33.$$

As another example, consider the poster design mentioned in the second paper.[11] In *D* four colors are used—*BG* 5/4, *R* 4/2, *R* 4/6, *BG* 7/4. Thus,

Number of colors	= 4
Pairs differing in hue	= 4
Pairs differing in value	= 5
Pairs differing in chroma	= 5
	$C=18$.

It has been shown[11] that the scalar moments are all equal, so $A=C_2^4=6$.

The usual table is then constructed (Table XIV). A consideration of all the combinations of hues for the four colors shows that the number of pairs with identical hues is 2 while the number with contrasting hues (Fig. 1) is 4. Value and chroma are considered with the help of Fig. 4 of the second paper.[11] For the pair *BG* 5/4, *R* 4/2

TABLE XIII.

	Identity	1st ambiguity	Similarity	2nd ambiguity	Contrast	Glare
Hue	X					
Value			X			
Chroma	X					

TABLE XIV.

	Identity	1st ambiguity	Similarity	2nd ambiguity	Contrast	Glare
Hue	XX				XXXX	
Value	X		XX	X	XX	
Chroma	X		XXX		XX	

the value difference is 1 and the chroma difference is 2, which gives a point in the region of similarity. Thus a check mark is made in the similarity column for both V and C. For BG 5/4, BG 7/4 the value difference is 2 and the chroma difference zero. Thus a mark is made to indicate identity in chroma and second ambiguity in value. The complete tabulation is shown above. The order is

$$O=A+2(1.5)+4(1.7)-1.3+2(0.7)-0.2 \\ +2(3.7)+0.8+3(0.1)+2(0.40) \\ =A+19.0=25.0.$$

The aesthetic measure of the design is therefore

$$M=25/18=1.39.$$

A word must be said about the inclusion of grays in a design. *When a gray is compared with a chromatic color it is considered to give no hue contribution, either to complexity or to order.* As an example, consider design B. The colors[11] are R 6/2, R 3/2, R 3/10, N 9. Thus,

Number of colors	= 4
Pairs differing in hue	= 0
Pairs differing in value	= 5
Pairs differing in chroma	= 5
	$C=14.$

As before, $A=6$. The hue relations are obtained by ignoring the achromatic element which gives only 3 hue pairs instead of 6. These 3 pairs are identities. Chroma and value relations are obtained exactly as with previous designs, giving Table XV. The order is

$$O=A+3(1.5)-1.3+0.8+5(3.7)+5(0.4) \\ =A+24.5=30.5,$$

TABLE XV.

	Identity	1st ambiguity	Similarity	2nd ambiguity	Contrast	Glare
Hue	XXX					
Value	X				XXXXX	
Chroma	X				XXXXX	

and the aesthetic measure is

$$M=30.5/14=2.18.$$

The foregoing examples all indicate designs of high aesthetic merit. In fact, any design having $M>0.5$ is desirable on the basis of abstract aesthetics, though the higher the aesthetic measure the better is the color harmony. Table XVI gives an indication of the range of M to be ex-

TABLE XVI. Calculated aesthetic measures.

Class of harmony		Colors	C	O	M Without area balance	M With area balance
One variable						
I, 1$A(a)$	(1)	N 7, N 8	3	2.7	0.90	1.23
	(2)	N 3, N 6	3	5.7	1.90	2.23
	(b)(1)	N 6, 7, 8	6	4.2	0.70	1.20
	(2)	N 3, 6, 9	6	14.1	2.35	2.85
	(3)	N 2, 3, 7	6	11.1	1.85	2.35
1$B(a)$	(1)	R 6/8, 5/8	3	3.0	1.00	1.33
	(2)	P 6/8, 3/8	3	6.0	2.00	2.33
	(b)(1)	GY 7/8, 6/8, 5/8	6	8.1	1.35	1.85
	(2)	R 8/4, 5/4, 2/4	6	18.0	3.00	3.50
	(3)	BG 7/4, 6/4, 3/4	6	12.2	2.03	2.53
2 (a)	(1)	YR 4/4, 4/8	3	0.3	0.10	0.43
	(2)	GY 7/2, 7/10	3	0.6	0.20	0.53
	(b)(1)	RP 4/4, 4/8, 4/12	6	1.2	0.20	0.70
	(2)	N 4, R 4/7, R 4/14	6	−1.2	−0.20	+0.30
3 (a)	(1)	Y 8/8, YR 8/8	3	0.6	0.20	0.53
	(2)	5R 7/8, 10GY 7/8	3	1.2	0.40	0.73
	(b)(1)	RP 3/10, P 3/10, PB 3/10	6	1.35	0.23	0.73
	(2)	G 7/2, YR 7/2, P 7/2	6	3.6	0.60	1.10
	(3)	GY 6/6, G 6/6, P 6/6	6	3.0	0.50	1.00
Two variables						
II, 1$A(a)$	(1)	BG 5/6, 4/4	4	2.3	0.58	0.83
	(2)	GY 7/4, 5/10	4	5.6	1.40	1.65
	(b)(1)	G 5/2, 6/4, 7/6	9	5.9	0.66	0.99
	(2)	R 7/6, R 4/4, R 1/2	9	16.8	1.87	2.20
	(3)	GY 5/2, 7/8, 8/10	9	13.5	1.50	1.83
1$B(a)$	(1)	R 5/6, 4/4, 4/8	8	6.9	0.87	1.24
	(2)	PB 7/6, 4/2, 4/10	8	16.8	2.10	2.48
	(3)	YR 7/4, 4/2, 4/6	8	13.5	1.69	2.06
	(b)(1)	N 5.5, G 6/2, G 5/2	8	3.9	0.49	0.86
	(2)	RP 5/2, 2/4, 8/4	8	16.8	2.10	2.48
	(3)	N 4.5, BG 5/8, BG 4/8	8	3.9	0.49	0.86
	(c)(1)	BG 6/6, 5/4, 5/6	7	4.9	0.70	1.13
	(2)	Y 8/4, 8/12, 5/4	7	12.2	1.74	2.17
	(3)	B 7/2, 7/6, 3/6	7	11.9	1.70	2.12
1$C(a)$		YR 7/6, 7/10, 6/6, 6/10	12	9.2	0.77	1.27
	(b)	N 6, B 6/8, N 3, B 3/8	12	17.9	1.49	1.99
	(c)	B 7/2, 7/6, 4/2, 4/6	12	16.2	1.35	1.85
1D		R 6/6, 4/2, 4/10, 2/6	16	11.8	0.74	1.11
2	(a)	RP 7/4, R 7/8	4	−0.1	−0.02	+0.22
	(b)	N 7, BG 7/4, G 7/4	6	−1.8	−0.30	+0.20
	(c)	G 5/4, BG 5/4, N 5, RP 5/4, R 5/4	14	+1.2	+0.09	+0.80
	(d)	YR 6/6, RP 6/6, PB 6/6, BG 6/6, GY 6/6	15	6.8	0.46	1.12
	(e)	R 5/6, P 5/4, BG 5/2, Y 5/4	15	−1.3	−0.09	+0.31
3 (a)	(1)	GY 7/6, G 6/6	4	2.6	0.65	0.90
	(2)	G 3/2, P 7/2	4	6.2	1.55	1.80
	(3)	R 7/8, RP 3/8	4	5.6	1.40	1.65
	(4)	YR 7/4, BG 6/4	4	3.2	0.80	1.05
	(b)(1)	R 2/4, G 7/4, B 7/4	5	12.5	2.51	3.11
	(2)	GY 8/6, R 4/6, PB 4/6, P 2/6	9	22.1	2.48	3.14
Three variables						
III, 1(a)		P 6/8, P 3/8, RP 6/8	7	12.2	1.75	2.16
	(b)	R 8/4, R 5/4, R 2/4, BG 5/4	12	24.2	2.02	2.52
	(c)	P 6/8, P 3/8, B 6/8, B 3/8	12	22.6	1.89	2.48
	(d)	N 4, R 7/6, R 6/4, B 7/6, G 7/6	25	11.5	0.46	0.86
2(a)		N 6, R 6/4, BG 6/4, BG 5/4	12	5.8	0.48	0.98
	(b)	RP 7/4, R 7/8, R 3/8	9	11.1	1.23	1.57
	(c)	BG 6/2, R 6/10, BG 2/2, R 2/10	16	25.2	1.57	1.95
	(d)	RP 3/2, R 3/10, RP 7/2, RP 6/2	15	25.6	1.70	2.10
3(a)		GY 8/6, R 4/6, PB 4/6, P 2/6, R 4/10	26	45.0	1.73	2.12
	(b)	N 5.5, B 6/4, R 5/4, R 5/12	16	9.3	0.58	0.96
	(c)	G 3/2, P 7/2, G 3/6, P 7/6	16	24.6	1.54	1.91
	(d)	R 7/8, RP 3/8, R 7/4, N 7	14	13.1	0.94	1.36
4(a)		N 5, GY 2/2, PB 8/2	9	14.4	1.60	1.93
	(b)	P 3/10, PB 4/8, RP 4/8, GY 7/4	20	21.3	1.07	1.57

pected in designs worked out by the methods of the previous papers. The classification of harmonies in the table is the same as in the first paper.[10] A study of these ratings leads to interesting conclusions, among which may be mentioned:

Correctly related grays give essentially as high an aesthetic measure as chromatic colors (see Class I, 1*A*.)

A harmony with constant hue tends to be very pleasing (for instance, Class I, 1*B*.)

Constant value, on the other hand, tends to give low aesthetic measures (as in I, 2 or II, 2.)

Simple designs with constant hue and constant chroma (I, 1*B*) are often more satisfactory than very complicated combinations with many hues (III, 1(*d*).)

Any of the examples of Table XVI or of the previous tables can be materialized very easily by means of Munsell papers or Allcolor papers. Thus anyone can try these various harmonies and see for himself how satisfactory the *M* rating is.

7. SUMMARY

Classical color harmony divides color combinations into two classes which may be called harmonies and disharmonies. As pointed out previously,[10] however, there are no such distinct classes but color combinations can be rated on a continuous scale from very poor combinations to very good ones. The method of aesthetic measure developed in the present paper provides a numerical rating scheme and gives a quantitative evaluation of the aesthetic merit of any given combination of colors.

According to Birkhoff, aesthetic measure is always equal to the number of elements of order divided by the complexity. The application of this general formula to color requires the determination of the weights to be given to elements of order and complexity. Such factors as identity, similarity, contrast, and ambiguity were evaluated by means of experiments made with color chips. Area balance was evaluated similarly. The results of all experiments are summarized in Table XI, and instructions are given for the calculation of the aesthetic measure of any color combination.

For convenience, the present paper has utilized experiments made with Munsell color chips and with the eyes adapted to a neutral gray. The method of aesthetic measure, however, is not limited to such colors or to this particular adaptation. It applies to any kind of color, including colored lights, and to any state of visual adaptation. The general case of adaptation will be treated in another paper.

The subject of aesthetic measure offers excellent possibilities for further research. The desirability of a more precise determination of the weights assigned to the various elements of order has already been mentioned. Also the method ignores a number of factors which may be important. For example, the hue preference of the individual[12] is not taken into account, all hues being weighted equally in our system. Two designs may have equal aesthetic measures but one may be much more suitable for a certain application than another because of its more appropriate range of chromas and values. Quiet designs are better in some cases than designs of greater contrast, yet the aesthetic measure may be less for the former than for the latter. There is also the question of whether people with different degrees of color education will come to the same conclusion in rating a harmony. Birkhoff assumed that there was such a unique aesthetic measure and we have followed his example. But perhaps a study of aesthetic sophistication would be desirable. The extension of the methods of this paper to the more general subject of pictorial composition would be another field for research.

[12] J. P. Guilford, J. Opt. Soc. Am. **30**, 455 (1940).

Proposed American Standard Practice for Microfilms—Z38.7.8*

1. DEFINITIONS

1.1 Microfilm shall mean a transparent flexible material carrying microphotographs for optical, but not cinematographic, projection or viewing.

1.2 A microphotograph is a reduced-size photographic documentary reproduction, generally too small to be read by the unaided eye. The photographs are made usually for record purposes from documentary material, such as texts or drawings, or from physical objects.

2. MATERIAL

2.1 Microfilms shall be made only on safety (slow-burning) stock as defined by American Standard Definition of Safety Photographic Film, Z38.3.1–1943, or the latest revision thereof approved by the American Standards Association. Microfilms intended for permanent preservation should be made on film stock meeting the specifications of the National Bureau of Standards for permanent record films as given in National Bureau of Standards Miscellaneous Publication M 158.

2.2 Thirty-five- and 16-mm stock used for microfilms shall conform in width and perforations, if latter are present, with American Standards for Cutting and Perforating Negative and Positive Raw Stock: 16-mm Sound Film, Z22.12–1941; 35-mm Film, Z22.34–1930; and 16-mm Silent Film, Z22.5–1941; or the latest revisions thereof approved by the American Standards Association. Double or single perforated or unperforated film may be used.

2.3 Microfilms normally are wound with the emulsion out for negatives and with emulsion in for positive prints.

3. FORMAT AND PLACEMENT

3.1 The width of the image on 35-mm film shall not exceed 24 mm (0.946 inch) on double perforated film; 28.6 mm (1.126 inches) on single perforated film; and 31.75 mm (1.250 inches) on unperforated film.

3.2 The width of the image on 16-mm film shall not exceed 10.4 mm (0.410 inch) on double perforated film; 12.7 mm (0.500 inch) on single perforated film; and 14.99 mm (0.590 inch) on unperforated film.

3.3 When framing in the reader is important, successive images should be separated by at least 0.5 mm (0.02 inch), and on strips of 10 feet (3 m) or longer, the separation should approach this value to conserve film and storage space.

3.4 The images on the film for library use should be arranged, when possible, as shown in the accompanying illustration. I. Single page of copy arranged lengthwise of the film. II. Copy arranged parallel to the sides of the film: IIA double-page; IIB single-page. III. Double-page copy arranged lengthwise of the film.

3.5 Reduction is the ratio of a linear dimension of the object to the corresponding dimension of the image on the film.

4. LEADER AND TRAILER

4.1 When the film is long enough to be stored on reels, a leader and a trailer of 18 inches (45 cm) each should be provided.

5. TITLE FRAME

5.1 When the film is stored on reels, the first image should be the title, or a condensed version of the title, in letters at least 0.08 inch (2 mm) high. The last frame should indicate the end of the film by "End" or other suitable designation.

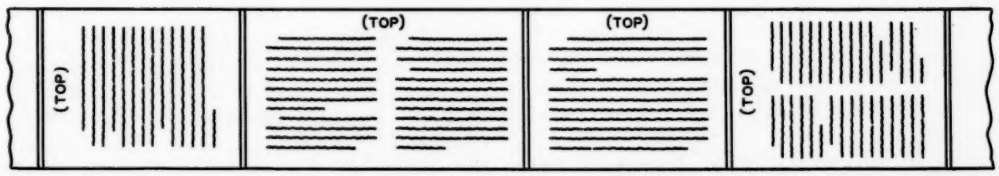
6. REELS

6.1 Microfilms 25 feet (7.6 m) long or longer should be wound on standard reels as defined herein.

6.2 Reels for holding microfilm should be of 50- or 100-foot capacity.

6.3 One-hundred-foot reels for both 35-mm and 16-mm microfilm should have an outside diameter of 3.678 ± 0.015 inches (93.42 ± 0.38 mm); a core of 1.25 ± 0.015 inches (31.75 ± 0.38 mm); one flange shall have a square hole of 0.320 ± 0.002 inch (8.13 ± 0.05 mm), and the other a similar square hole or a round hole of 0.320 ± 0.002 inch (8.13 ± 0.05 mm) in diameter.

6.3.1 Fifty-foot reels should have an outside diameter of 2.906 ± 0.015 inches (73.8 ± 0.38 mm), with other dimensions the same as given for 100-foot reels.



* Not Approved. Published for Trial and Criticism January 1, 1944 by American Standards Association. Sponsor: Optical Society of America.

Proposed American Standard Specifications for Microfilm Readers—Z38.7.9*

1. DEFINITION

1.1 A microfilm reader is defined as a projection device for showing a readable image of a microfilm (cf. Proposed American Standard Practice for Microfilms, Z38.7.8), usually on a self-contained screen.

2. DIMENSIONS

2.1 Readers for 35-mm film shall accommodate film 35 mm +0.05, −0.75 mm (1.378 inches +0.002, −0.030 inch) wide; and for 16-mm film, 16 mm +0.02, −0.37 mm (0.630 inch +0.001, −0.015 inch) wide.

2.2 The aperture of the film gate should include at least 24 by 24 mm (0.946 by 0.946 inch) when designed for 35-mm double perforate film, and 12.7 by 12.7 mm (0.5 by 0.5 inch) when designed for 16-mm double perforate film. When the reader is designed for unperforated 35-mm film, the aperture should accommodate a film image 31.75 mm (1.250 inches) wide; and for 16-mm unperforated film, 14.99 mm (0.590 inch) wide.

2.3 In readers for 35-mm film, there should be provision for moving the film across the aperture for viewing images of varying length, and for moving the film laterally so that images up to 31.75 mm (1.25 inches) wide may be projected, although the whole width need not be projected at the same time.

2.4 Readers for library use should have provision for rotating the film or the image beam, so that images taken as described in Z38.7.8, Paragraph 3.4, may be shown right side up on the screen.

2.5 The image should be properly oriented on the screen when negative film is wound emulsion side out, and positive film, emulsion side in.

2.6 Means should be provided to release pressure on the image area during movement of the film and to hold the film flat during use.

3. LENS

3.1 The image should not require refocusing as a line of print is read from one side across the center to the opposite side.

3.2 Color or distortion, or both, should not be obvious in the screen image of black and white printed matter.

4. MAGNIFICATION

4.1 Readers with fixed (not adjustable) magnification ordinarily should provide a minimum magnification of 12× for 35-mm film and 22× for 16-mm film.

4.2 Magnification shall be defined as the ratio of a linear dimension of the screen image to the length of the same object on the film and shall be expressed as diameters (×).

5. ILLUMINATION

5.1 The ratio of corner-to-center illumination of the screen should be not less than 65 percent, and any two adjacent areas shall not vary from each other by more than 20 percent, when the illumination is measured as specified in American Standard Methods of Testing Printing and Projection Equipment, Z38.7.5, Section 3.

Corner illumination shall be measured as close to the extreme corners of the field as possible. If the measuring instrument integrates the illumination over an appreciable area, the maximum dimension of that area shall not exceed one-fifth of the width of the illuminated area.

5.2 There should be no direct or reflected light distracting to the operator. No direct light should pass into the room, and the leakage of stray light should be avoided insofar as practical.

6. TEMPERATURE

6.1 The temperature of the film should not exceed 75 C (167 F) when measured by the testing method specified in Z38.7.5, Section 2.

7. REELS

7.1 When provision is made for projecting film on reels, standard reels—as defined in Z38.7.8—shall be accommodated.

8. SPROCKETS

8.1 If sprockets are provided for moving the film, they should conform to American Standards for 35-mm and 16-mm film, respectively (35-mm Film: 16-Tooth Projector Sprockets, Z22.35–1930; 16-mm Film: Projector Sprockets, Z22.6–1941), or the latest revision thereof approved by the American Standards Association.

* Not Approved. Published for Trial and Criticism January 1, 1944 by American Standards Association. Sponsor: Optical Society of America.